A Nonparametric Measure of Heteroskedasticity *

Xiaojun Song† Abderrahim Taamouti‡
Peking University Durham University

June 17, 2020

ABSTRACT

We introduce a nonparametric measure to quantify the degree of heteroskedasticity at a fixed quantile of the conditional distribution of a random variable. Our measure of heteroskedasticity is based on nonparametric quantile regressions and is expressed in terms of unrestricted and restricted expectations of quantile loss functions. It can be consistently estimated by replacing the unknown expectations by their nonparametric estimates. We derive a Bahadur-type representation for the nonparametric estimator of the measure. We provide the asymptotic distribution of this estimator, which one can use to build tests for the statistical significance of the measure. Thereafter, we establish the validity of a fixed regressor bootstrap that one can use in finite-sample settings to perform tests. A Monte Carlo simulation study reveals that the bootstrap-based test has a good finite sample size and power for a variety of data generating processes and different sample sizes. Finally, two empirical applications are provided to illustrate the importance of the proposed measure.

Keywords: Bootstrap; measuring heteroskedasticity; nonparametric quantile regressions; income and years of education.

Journal of Economic Literature classification: C12; C14; C15; C19; G1; G12; E3; E4.

---

*The authors thank the Executive Editors and two anonymous referees for several useful comments. Financial support from China's National Key Research Special Program Grants 2016YFC0207705, the National Natural Science Foundation of China (Grant No. 71532001, 71973005), and Key Laboratory of Mathematical Economics and Quantitative Finance (Peking University), Ministry of Education, China, is acknowledged.

† *Corresponding author:* Department of Business Statistics and Econometrics, Guanghua School of Management and Center for Statistical Science, Peking University, Beijing, 100871, China. E-mail: sxj@gsm.pku.edu.cn.

‡ Department of Economics and Finance, Durham University Business School. Address: Mill Hill Lane, Durham, DH1 3LB, UK. Tel: +44-1913345423. E-mail: abderrahim.taamouti@durham.ac.uk.

1 Introduction

Regression errors in cross–section and time series models frequently exhibit heteroskedasticity. Even though the latter was always viewed as a problem that one needs to treat to improve efficiency, some authors take a different view and argue that the heterogeneity in the degree of heteroskedasticity can often have important theoretical and substantive implications over and above those for accurate inference. Among others, Newey and Powell (1987) argue that the change in the degree of heteroskedasticity in the conditional distribution of the dependent variable might be viewed as a symptom of misspecification of the regression function [e.g. indicates the presence of an omitted variable].[1] In addition, Arabmazar and Schmidt (1981) study the impact of the degree of heteroskedasticity in the error terms on the size of the inconsistency of the MLE estimator. They show that the inconsistency is greater the greater the degree of heteroskedasticity. Others have stressed the importance of understanding the economic meaning of heteroskedasticity when its degree changes across the conditioning variables. For example, Meghir and Pistaferri (2004) among others point out the relevance of allowing for variance of income to vary across different education levels for modelling earnings distribution. Much research has been devoted to building tests of heteroskedasticity. However, to the best of our knowledge, no research really develops measures of the degree of heteroskedasticity. The present paper aims to propose a nonparametric measure of the degree of heteroskedasticity at a given fixed quantile of the conditional distribution of a random variable. The measure can also be used to test for heteroskedasticity.

Measuring the degree of heteroskedasticity might also help us to better understanding the relationship between the latter and the efficiency of the estimates of regression coefficients. The presence of a *strong* heteroskedasticity in the data negatively affects the performance (size and power) of classical tests such as t–test and F–test. Several econometric textbooks and papers have reported that the available heteroskedasticity consistent standard errors lead to tests/confidence intervals that deliver substantial under or over size/coverage depending on the degree of heteroskedasticity; see Kennedy (1985), Dufour (2003), Cribari-Neto (2004), Dufour and Taamouti (2010), Hausman and Palmer (2012), Cattaneo, Jansson, and Newey (2018), among others. For a valid inference, different estimation techniques/tests might need to be applied depending on the degree of heteroskedasticity. Senyo and Adjibolosoo (1989) argue that if the degree of heteroskedasticity is not high, then the ordinary least squares (OLS) estimator might still perform better than the generalized least squares (GLS) estimator. They stress the importance of developing measures of the strength of heteroskedasticity, which can help researchers understand when to use either the OLS estimator or the GLS estimator.

The above motivations illustrate the usefulness of providing measures of the degree of heteroskedas-

[1] Downs and Rocke (1979) provide some real examples that show how the degree of heteroskedasticity indicates that other variables other than the ones considered in the analysis are needed for modelling the dependent variable of interest.

ticity in the conditional distribution of the dependent variable. In this paper, we introduce a measure of heteroskedasticity using nonparametric quantile regressions. This measure can quantify the degree of heteroskedasticity at a fixed quantile of the conditional distribution of the variable of interest. Unfortunately, the existing heteroskedasticity tests fail to accomplish this task, because they only provide evidence on the presence or the absence of heteroskedasticity, and statistical significance depends on the available data and test power. A strong heteroskedasticity may not be statistically significant, and a statistically significant heteroskedasticity may not be strong from an economic viewpoint. To the best of our knowledge, this is the first measure of heteroskedasticity, which is based on nonparametric quantile regressions and expressed in terms of unrestricted and restricted expectations of quantile loss functions. It is consistently estimated by replacing the unknown expectations by their nonparametric estimates. Our theoretical results are proven under the assumptions of dependent data, but they are still valid for cross-sectional data.

We also note that there is an abundant literature on nonparametric quantile regression when parametric quantile regression function is not available. For example, Chaudhuri (1991), Yu and Jones (1998) and Guerre and Sabbah (2012) consider nonparametric estimation of conditional quantiles for i.i.d. observations by using local polynomial regression, while Honda (2000), Hall et al. (2002) and Kong et al. (2010) examine the asymptotic properties of the estimator of Chaudhuri (1991) for strongly mixing stationary processes. Nevertheless, none of the aforementioned estimators is designed to measure the degree of heteroskedasticity of the conditional distribution of a random variable. Our paper fills this gap by suggesting a convenient R^2–type measure of heteroskedasticity at a fixed quantile based on the nonparametric quantile estimators.

Furthermore, we derive a Bahadur–type representation for the nonparametric estimator of the measure. We provide its asymptotic distribution, which one can use to build tests for the statistical significance of the measure. Moreover, since testing that the value of the measure is equal to zero is equivalent to testing for homoscedasticity, another contribution of this paper consists in providing a test for heteroskedasticity that is robust to the parametric misspecification of conditional location function. The existing parametric specification-based tests for heteroskedasticity generally suffer from the model misspecification problem, and require correct parametric specification of the regression function. Thereafter, we establish the validity of a fixed regressor bootstrap that one can use in finite–sample settings to perform tests for different values of the measure. A Monte Carlo simulation study reveals that the bootstrap test has a good finite sample size and power for a variety of data generating processes and different sample sizes.

Two empirical applications are also provided to illustrate the importance of the proposed measure. In the first application, we are interested in measuring the degree of heteroskedasticity of income conditional on the years of education, and in the second application, we quantify the degree of heteroskedasticity for 30 stock returns. For the first application, our results show that the degree of income variation decreases when the years of education increase. Thus, the income of highly educated people varies less compared with the income of those with low levels of education. Furthermore, we find that the degree of income variation for females is generally smaller than the degree of income variation for males. For the second application, the results confirm that all stock returns under consideration are conditionally heteroskedastic. In addition, these results show that there is some heterogeneity in the degree of heteroskedasticity across the stocks.

To sum up, our contributions are threefold. Firstly, we propose a fully model–free measure to gauge the degree of heteroskedasticity. Secondly, we show that the proposed measure can be used as a test to detect heteroskedasticity. Our test is designed to be particularly robust to the misspecification in the conditional mean and is able to capture various forms of conditional heteroskedasticity. Lastly, we propose an innovative bootstrap procedure to implement the test.

This paper is organized in the following way. Section 2 presents the general theoretical framework that underlies the definition of the measure of heteroskedasticity in the presence of constant and non-constant means. In Section 3, we discuss the estimation of nonparametric quantile regressions and, consequently, of the measure of heteroskedasticity based on the local polynomial estimation of the unrestricted and restricted expectations of quantile loss functions. We derive a Bahadur-type representation for the nonparametric estimator of the measure. We also provide the asymptotic distribution of this estimator, which one can use to build tests for the statistical significance of the measure. In Section 4, we establish the validity of a fixed regressor bootstrap that one can use in finite-sample settings to perform tests. Section 5 presents a Monte Carlo simulation exercise to investigate the finite sample properties of the bootstrap-based test of the measure of heteroskedasticity. Section 6 is devoted to an empirical application, and the conclusion relating to the results is given in Section 7. The main assumptions of the paper and the proofs of the theoretical results are presented in the appendices A.2.1 and A.2.2, respectively.

2 Framework

Let $\left\{(Y_t, X_t) \in \mathbb{R} \times \mathbb{R}^d \equiv \mathbb{R}^{d+1},\ t \in \mathbb{Z}\right\}$ be a strictly stationary stochastic process in $\mathbb{R}^{d+1}$ for a fixed known integer $d \geq 1$. We are interested in the conditional variance of Y_t conditional on X_t, and we consider the following nonparametric mean regression:

$$Y_t = m\left(X_t\right) + \sigma\left(X_t\right)\epsilon_t, \tag{1}$$

where $m\left(X_t\right)$ and $\sigma\left(X_t\right) > 0$ are some smooth and unknown functions for the conditional location (mean) and the conditional scale (standard deviation), respectively, and ϵ_t is an error term independent of X_t and its past. The conditional quantile functions of Y_t conditional on X_t are then simply

$$Q^{(\tau)}\left(Y_t \middle| X_t\right) = m\left(X_t\right) + \sigma\left(X_t\right) D^{-1}\left(\tau\right),\ \text{for } \tau \in (0,1)\,, \tag{2}$$

where $D\left(\cdot\right)$ is the cumulative distribution function (CDF) of the error term ϵ.

In the next sections, we provide measures of the degree of heteroskedasticity at a fixed quantile of the conditional distribution of Y_t conditional on X_t and across different ranges of the conditioning variable X_t. In particular, we show how to measure the heteroskedasticity under different scenarios concerning the conditional mean of Y_t. We distinguish between the two scenarios: (1) the mean function $m(X_t)$ is constant and (2) the mean function $m(X_t)$ is not constant. Furthermore, we show that these measures can be used to build tests of homoscedasticity. Henceforth, whenever homoscedasticity holds, we shall use $Var(Y_t|X_t) = \sigma^2(X_t) = \sigma_0^2$ almost surely (a.s.) to denote this case.

2.1 Constant mean

Here, we assume that $m(X_t) = \mu$, where μ is an unknown constant that is equal to the unconditional mean of Y_t. Under this assumption and using Equation (2), the conditional τ-th quantile of Y_t given X_t becomes

$$Q^{(\tau)}(Y_t|X_t) = \mu + \underbrace{\sigma(X_t) D^{-1}(\tau)}_{=\phi(X_t,\tau)} = \mu + \phi(X_t,\tau), \text{ for } \tau \in (0,1). \tag{3}$$

From Equation (3), testing for homoscedasticity is equivalent to testing that the quantile function $Q^{(\tau)}(Y_t|X_t)$ does not depend on X_t. In other words, the null hypothesis of homoscedasticity and the alternative hypotheses of heteroskedasticity can be expressed as follows:

$$\left[\begin{array}{l} H_0 : Q^{(\tau)}(Y_t|X_t) = \mu + \sigma_0 D^{-1}(\tau) = \mu + \xi(\tau) \\ \text{vs} \\ H_A : Q^{(\tau)}(Y_t|X_t) = \mu + \sigma(X_t) D^{-1}(\tau) = \mu + \phi(X_t,\tau). \end{array}\right. \tag{4}$$

Observe that the hypothesis testing problem in (4) is equivalent to the problem of assessing model adequacy in quantile regressions; see for example Noh et al. (2013). Thus, measures similar to the well-known coefficient of determination R^2, but for quantile regressions, can be used to measure/test heteroskedasticity. Hence, to quantify the degree of heteroskedasticity at a given fixed quantile, we propose the following measure, which we express in terms of unrestricted and restricted expectations of quantile loss functions

$$H(\tau) = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \sigma(X_t) D^{-1}(\tau)\right)\right]}{E\left[\rho_\tau\left(Y_t - \mu - \sigma_0 D^{-1}(\tau)\right)\right]} = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \phi(X_t,\tau)\right)\right]}{E\left[\rho_\tau\left(Y_t - \mu - \xi(\tau)\right)\right]}, \tag{5}$$

where the check loss function $\rho_\tau(\cdot)$ is defined as follows:

$$\rho_\tau(u) \equiv (\tau - 0.5)u + 0.5|u| = u(\tau - \mathbb{I}(u < 0)),$$

with $\mathbb{I}(u < 0)$ as an indicator function that takes the value 1 when $u < 0$ and the value 0 when $u \geq 0$; see Koenker (2005). To the best of our knowledge, this is the first time that the measure $H(\tau)$ has been used to quantify heteroskedasticity. This measure, by construction, satisfies $0 \leq H(\tau) \leq 1$, as we have $0 \leq E\left[\rho_\tau\left(Y_t - \mu - \phi(X_t,\tau)\right)\right] \leq E\left[\rho_\tau\left(Y_t - \mu - \xi(\tau)\right)\right]$. In contrast, if $\phi(X_t,\tau)$ is constant for a given τ, then it can be seen that the measure $H(\tau)$ is equal to 0, which corresponds to the homoscedasticity case. On the other hand, $H(\tau) > 0$ corresponds to the heteroskedasticity case, and the larger the $H(\tau)$, the stronger the degree of heteroskedasticity in regression (1). Thus, like the classical coefficient of determination R^2, $H(\tau)$ can be understood as an index of homoscedasticity adequacy.

Moreover, notice that $H(\tau)$ can be used to build tests of homoscedasticity. Under the null of homoscedasticity, we have $\phi(X_t, \tau) = \xi(\tau)$ with probability one:

$$E\left[\rho_\tau\left(Y_t - \mu - \phi(X_t, \tau)\right)\right] = E\left[\rho_\tau\left(Y_t - \mu - \xi(\tau)\right)\right],$$

and thus, $H(\tau) = 0$. Hence, testing homoscedasticity in regression (1) is equivalent to testing $H(\tau) = 0$.

Finally, the above measure can generally be written as follows:

$$H^w(\tau) = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \phi(X_t, \tau)\right) w(X_t)\right]}{E\left[\rho_\tau\left(Y_t - \mu - \xi(\tau)\right) w(X_t)\right]}, \tag{6}$$

where $w(\cdot)$ is a non-negative weighting function that is continuous on a compact support. An interesting example of the weighting function $w(\cdot)$ is given by the following:

$$w(X_t) = \mathbb{I}\left[X_t \leq q^x(\alpha)\right] = \begin{cases} 1 \text{ if } X_t \leq q^x(\alpha) \\ 0, \text{ otherwise,} \end{cases} \tag{7}$$

where $q^x(\alpha)$ is the α-th quantile of X_t for $\alpha \in (0,1)$. By selecting $w(\cdot)$ as in (7), we can measure the degree of heteroskedasticity of Y when X takes values below its unconditional lower quantile $q^x(\alpha)$. Similarly, a measure of the degree of heteroskedasticity of Y when X takes values above its unconditional upper quantile $q^x(1-\alpha)$ is given by the following:

$$H^w(\tau) = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \phi(X_t, \tau)\right) \mathbb{I}\left[X_t \geq q^x(1-\alpha)\right]\right]}{E\left[\rho_\tau\left(Y_t - \mu - \xi(\tau)\right) \mathbb{I}\left[X_t \geq q^x(1-\alpha)\right]\right]}, \tag{8}$$

where $\mathbb{I}\left[X_t \geq q^x(1-\alpha)\right]$ is an indicator function that takes the value 1 if $X_t \geq q^x(1-\alpha)$ and 0 otherwise.

2.2 Non-constant mean

In the previous subsection, we assume that the conditional mean of Y_t is constant ($m(X_t) = \mu$). However, if $m(X_t)$ is not constant, then the above procedure for measuring heteroskedasticity is useless because both under the null and alternative $Q^{(\tau)}(Y_t | X_t)$ will be a function of X_t, except if the conditional τ-th quantile of Y_t is zero. To overcome this situation, we propose the following two-stage procedure. Unlike the constant mean case, we first need to filter out the effect of X_t on the mean of Y_t. To this end, we consider the transformed variable $\overline{Y}_t = Y_t - m(X_t)$, and the nonparametric regression in (1) becomes

$$\overline{Y}_t = \sigma(X_t)\, \epsilon_t.$$

In practice, however, the functional form of $m(X_t)$ is unknown. In the next section, we discuss the estimation of $m(X_t)$ and its effect on the asymptotic properties of the estimated measure of heteroskedasticity.

Next, under the assumption that ϵ_t is independent of X_t, the conditional τ-th quantile of $\overline{Y}_t$ conditional on X_t is given by the following:

$$\overline{Q}^{(\tau)}\left(\overline{Y}_t \middle| X_t\right) = \sigma(X_t) D^{-1}(\tau) = \overline{\phi}(X_t,\tau), \text{ for } \tau \in (0,1). \tag{9}$$

From Equation (9), testing for the homoscedasticity in regression (1) is equivalent to testing that the quantile function $\overline{Q}^{(\tau)}\left(\overline{Y}_t \middle| X_t\right)$ does not depend on X_t. In other words, the null of homoscedasticity and the alternative of heteroskedasticity for a given fixed quantile can be expressed as follows:

$$\left[\begin{array}{l} H_0 : \overline{Q}^{(\tau)}\left(\overline{Y}_t \middle| X_t\right) = \overline{\xi}(\tau) \\ \text{vs} \\ H_A : \overline{Q}^{(\tau)}\left(\overline{Y}_t \middle| X_t\right) = \overline{\phi}(X_t,\tau). \end{array}\right. \tag{10}$$

The testing problem in (10) is equivalent to the problem of assessing model adequacy in a nonparametric quantile regression framework with the mere exception of the dependent variable now being $\overline{Y}_t := Y_t - m(X_t)$. Thus, to quantify/test heteroskedasticity at a given fixed quantile, we propose the following measure that is based on nonparametric quantile regressions and expressed in terms of unrestricted and restricted expectations of quantile loss functions:

$$H(m,\tau) = 1 - \frac{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\phi}(X_t,\tau)\right)\right]}{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\xi}(\tau)\right)\right]}.$$

Observe that the measure $H(m,\tau)$ satisfies $0 \le H(m,\tau) \le 1$. On the one hand, if $\overline{\phi}(X_t,\tau)$ is constant for a given τ, then it can be seen that $H(m,\tau)$ is equal to 0, which corresponds to the homoscedasticity case. On the other hand, $H(m,\tau) > 0$ corresponds to the heteroskedasticity case, and the larger the $H(m,\tau)$, the stronger the degree of heteroskedasticity in regression (1).

Moreover, notice that $H(m,\tau)$ can be used to build tests of homoscedasticity. Under the null of homoscedasticity, we have $\overline{\phi}(X_t,\tau) = \overline{\xi}(\tau)$ with probability 1:

$$E\left[\rho_\tau\left(\overline{Y}_t - \overline{\phi}(X_t,\tau)\right)\right] = E\left[\rho_\tau\left(\overline{Y}_t - \overline{\xi}(\tau)\right)\right],$$

and thus, $H(m,\tau) = 0$. Hence, testing homoscedasticity in (1) is equivalent to testing $H(m,\tau) = 0$.

Finally, as we saw in the previous subsection, the above measure can generally be written as follows:

$$H^w(m,\tau) = 1 - \frac{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\phi}(X_t,\tau)\right) w(X_t)\right]}{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\xi}(\tau)\right) w(X_t)\right]},$$

where $w(.)$ is a non-negative weighting function that we define in (7). This measure allows us to quantify the degree of heteroskedasticity of Y when X takes values below its unconditional lower quantile $q^x(\alpha)$ for

$\alpha \in (0,1)$. Similarly, a measure of the degree of heteroskedasticity of Y when X takes values above its unconditional upper quantile $q^x(1-\alpha)$ is given by the following:

$$H^w(m,\tau) = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \phi\left(X_t,\tau\right)\right)\mathbb{I}\left[X_t \geq q^x\left(1-\alpha\right)\right]\right]}{E\left[\rho_\tau\left(Y_t - \mu - \xi\left(\tau\right)\right)\mathbb{I}\left[X_t \geq q^x\left(1-\alpha\right)\right]\right]},$$

where $\mathbb{I}\left[X_t \geq q^x(1-\alpha)\right]$ is an indicator function that takes the value 1 if $X_t \geq q^x(1-\alpha)$ and 0 otherwise.

3 Estimation and asymptotic distribution

In this section, we introduce a nonparametric estimator for the measure of heteroskedasticity at a given fixed quantile and we study its asymptotic properties in the presence of the constant and non-constant means of Y. As we have seen in Section 2, the measure of heteroskedasticity is expressed in terms of unrestricted and restricted expectations of quantile loss functions. It can be estimated by replacing the unknown expectations of check loss functions by their nonparametric estimates from a finite sample. To obtain these nonparametric estimates and due to its well-known advantages, we propose using the local polynomial approach as discussed in Fan and Gijbels (1996).

3.1 Constant mean

Let $\left\{(Y_t, X_t) \in \mathbb{R}\times\mathbb{R}^d \equiv \mathbb{R}^{d+1},\ t = 1,...,T\right\}$ be a sample of strictly stationary stochastic processes Y and X. This sample will be used to estimate the following measure of heteroskedasticity at a given fixed quantile τ:

$$H^w(\tau) = 1 - \frac{E\left[\rho_\tau\left(Y_t - \mu - \phi\left(X_t,\tau\right)\right)w\left(X_t\right)\right]}{E\left[\rho_\tau\left(Y_t - \mu - \xi\left(\tau\right)\right)w\left(X_t\right)\right]}, \text{ for } \tau \in (0,1).$$

$H^w(\tau)$ can be estimated by replacing $E\left[\rho_\tau\left(Y_t - \mu - \phi\left(X_t,\tau\right)\right)w\left(X_t\right)\right]$ and $E\left[\rho_\tau\left(Y_t - \mu - \xi\left(\tau\right)\right)w\left(X_t\right)\right]$ by their nonparametric estimates from a finite sample. To do this, we need to estimate the quantities $Y_t - \mu - \phi(X_t,\tau)$ and $Y_t - \mu - \xi(\tau)$ and replace the theoretical expectations by their empirical analogs. We also need to assume that the conditional quantile function $\phi(x,\tau)$ is continuously differentiable up to order $p+1$. We then estimate $\phi(x,\tau)$ using the following multivariate p-th order local polynomial approximation:

$$\phi(z,\tau) \approx \sum_{0\leq|\underline{r}|\leq p} \frac{1}{|\underline{r}|} D^{|\underline{r}|}\phi(x,\tau)(z-x)^{|\underline{r}|},$$

where, for any $\underline{r} = (r_1,\cdots,r_d)$, $|\underline{r}| = \sum_{i=1}^d r_i$, $\underline{r}! = r_1! \times \cdots r_d!$, and

$$D^{\underline{r}}\phi(x,\tau) = \frac{\partial^{\underline{r}}\phi(x,\tau)}{\partial x_1^{r_1}\cdots\partial x_d^{r_d}}, \quad \underline{x}^{\underline{r}} = x_1^{r_1}\times\cdots\times x_d^{r_d}, \quad \text{and} \quad \sum_{0\leq|\underline{r}|\leq p} = \sum_{j=0}^{p}\underbrace{\sum_{r_1=0}^{j}\cdots\sum_{r_d=0}^{j}}_{r_1+\cdots+r_d=j}.$$

Instead of estimating the scale $\sigma(x)$, we estimate $\phi(x,\tau) := \sigma(x)D^{-1}(\tau)$ as a whole. This means that the assumption that $D^{-1}(\tau)$ is known a priori is not required. Koenker and Zhao (1996) have also estimated

$\phi(x,\tau)$ as a whole in the context of ARCH models. Moreover, observe that if the innovations ϵ_t follow a symmetric distribution, then $\phi(x,0.5)$ will be equal to 0 for any x, and it will not be identifiable. Thus, as long as the τ-th quantile of ϵ_t satisfies $D^{-1}(\tau) \neq 0$, the use of $H^w(\tau)$ to measure heteroskedasticity is justified. For more details on local polynomial estimation, the reader can consult Fan and Gijbels (1996) and Ruppert and Wand (1994), among others.

We next demean the dependent variable Y_t using $\widehat{\widetilde{Y}}_t = Y_t - \widehat{\mu}$, where the sample mean of Y_t is given by $\widehat{\mu} = T^{-1}\sum_{t=1}^{T} Y_t$. The nonparametric estimator of our measure, say $\hat{H}^w(\tau)$, is easily constructed through the following three steps:

(i) Estimate the τ-th marginal quantile of $\widehat{\widetilde{Y}}_t$, which we denote by $\widehat{\xi}(\tau)$, through minimizing the empirical check loss function $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\widetilde{Y}}_t - \theta\right)$ with respect to θ. Then, calculate $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\widetilde{Y}}_t - \widehat{\xi}(\tau)\right)w(X_t)$.

(ii) Estimate the τ-th quantile regression of $\widehat{\widetilde{Y}}_t$ on X_t through minimizing

$$\frac{1}{T}\sum_{t=1}^{T} K_h(X_t - x)\,\rho_\tau\left(\widehat{\widetilde{Y}}_t - \sum_{0\le|\underline{r}|\le p}\beta_{\underline{r}}(X_t - x)^{\underline{r}}\right), \tag{11}$$

where $K_h(u) = h^{-d}K(u/h)$ is a kernel for a d-dimensional product kernel function $K:\mathbb{R}^d \to \mathbb{R}$, and $h \equiv h_T \in \mathbb{R}^+$ is the usual bandwidth parameter converging to 0 at a proper rate when T tends to infinity. Assumptions on the kernel function K and the bandwidth parameter h are discussed in Appendix A.2.1. Denote by $\widehat{\beta}_{\underline{r}}$, for $0 \le |\underline{r}| \le p$, the minimizer of the function in (11). Then, the p-th order local polynomial estimator of $\phi(x,\tau)$ is given by $\widehat{\phi}(x,\tau) = \widehat{\beta}_0$, where $\widehat{\beta}_0$ is the first component of the vector $\widehat{\beta}_{\underline{r}}$. We obtain $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\widetilde{Y}}_t - \widehat{\phi}_{-t}(X_t,\tau)\right)w(X_t)$, where $\widehat{\phi}_{-t}(X_t,\tau)$ denotes the leave-observation-t-out estimator for $\phi(X_t,\tau)$.

(iii) Finally, estimate the measure

$$\hat{H}^w(\tau) = 1 - \left(T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\widetilde{Y}}_t - \widehat{\phi}_{-t}(X_t,\tau)\right)w(X_t)\right) \Big/ \left(T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\widetilde{Y}}_t - \widehat{\xi}(\tau)\right)w(X_t)\right). \tag{12}$$

We are now ready to state two main results for the constant mean case. Their proofs are similar to, but much simpler than, those of Theorems 3 and 4 below; therefore, they are omitted. Henceforth, we focus on $\tau \in \mathcal{T}$ with $\mathcal{T} = [a_1, a_2]$ for $0 < a_1 < a_2 < 1$. The following theorem establishes an asymptotic Bahadur representation of the estimator $\hat{H}^w(\tau)$ in (12).

Theorem 1 *Assume that the mean function $m(X_t)$ in the regression* (1) *is constant and unknown. Suppose Assumptions **C.1-C.10** in Appendix A.2.1 hold. If, furthermore, $p > d/2 - 1$ and $h = O(T^{-\kappa})$, with $1/(2p+2+d) < \kappa < 1/(2d)$, then for a given $\tau \in \mathcal{T}$, we have*

$$\sqrt{T}\left(\hat{H}^w(\tau) - H^w(\tau)\right) = (1 - H^w(\tau))\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(e_t - u_t) + o_p(1),$$

where

$$\begin{aligned} e_t &= \frac{\rho_\tau\left(Y_t-\mu-\xi\left(\tau\right)\right)w\left(X_t\right)-E\left[\rho_\tau\left(Y_t-\mu-\xi\left(\tau\right)\right)w\left(X_t\right)\right]}{E\left[\rho_\tau\left(Y_t-\mu-\xi\left(\tau\right)\right)w\left(X_t\right)\right]}, \\ u_t &= \frac{\rho_\tau\left(Y_t-\mu-\phi\left(X_t,\tau\right)\right)w\left(X_t\right)-E\left[\rho_\tau\left(Y_t-\mu-\phi\left(X_t,\tau\right)\right)w\left(X_t\right)\right]}{E\left[\rho_\tau\left(Y_t-\mu-\phi\left(X_t,\tau\right)\right)w\left(X_t\right)\right]}. \end{aligned}$$

An immediate implication of Theorem 1 is that the nonparametric estimator $\hat{H}^w(\tau)$ in (12) is consistent. Furthermore, by applying the central limit theorem on the weakly dependent process $\{(e_t-u_t)\}$ (e.g., Gao, 2007), Theorem 1 can be used to construct confidence intervals for $H^w(\tau)$. However, when homoscedasticity holds ($H^w(\tau)=0$), we see from the Bahadur representation that the asymptotic variance of $\hat{H}^w(\tau)$ degenerates to zero. This implies that the asymptotic normality that we obtain when we use Theorem 1 and the CLT is also degenerated and meaningless. Thus, unlike the cases where the degree of $H^w(\tau)$ is important (i.e., the value of the measure is not zero and large), we should investigate the next leading term in the Bahadur expansion in order to get a non-degenerated distributional result. Using the standard theory for U–statistics, the next theorem provides the limiting distribution of $\hat{H}^w(\tau)$ when homoscedasticity holds in regression (1).

Theorem 2 *Assume that the mean function $m(X_t)$ in the regression (1) is constant and unknown. Suppose Assumptions **C.1-C.10** in Appendix A.2.1 hold. If, furthermore, $p>d/2-1$ and $h=O(T^{-\kappa})$, with $1/(2p+2+d)<\kappa<1/(2d)$, then under the null hypothesis of homoscedasticity, for a given $\tau\in\mathcal{T}$, we have*

$$Th^{d/2}\hat{H}^w(\tau)\to_d N\left(0,\sigma_{0\tau}^2\right),$$

with

$$\sigma_{0\tau}^2=\frac{2\tau^2\left(1-\tau\right)^2}{r_\tau^2}\int K^2\left(u\right)du\int\frac{w^2\left(x\right)}{f_{\varepsilon,X}^2\left(0,x\right)}f_X^2\left(x\right)dx,$$

where $r_\tau=E\left[\rho_\tau\left(Y_t-\mu-\phi\left(X_t,\tau\right)\right)w\left(X_t\right)\right]$, $f_{\varepsilon,X}(0,x)$ is the joint density of quantile error $\varepsilon_t=Y_t-\mu-\phi(X_t,\tau)$ and X_t evaluated at $\varepsilon_t=0$, and $f_X(x)$ is the marginal density of X_t.

A consistent estimator of $\sigma_{0\tau}^2$ is given by

$$\widehat{\sigma}_{0\tau}^2=\frac{2\tau^2\left(1-\tau\right)^2}{\widehat{r}_\tau^2}\frac{1}{T\left(T-1\right)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\frac{w^2\left(X_t\right)}{\widehat{f}_{\widehat{\varepsilon},X}^2\left(0,X_t\right)}\frac{1}{h^d}K^2\left(\frac{X_t-X_s}{h}\right),$$

where

$$\widehat{r}_\tau=\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t-\widehat{\mu}-\widehat{\phi}_{-t}(X_t,\tau)\right)w(X_t)$$

is a consistent estimator of r_τ and $\widehat{f}_{\widehat{\varepsilon},X}(0,x)$ is the kernel density estimator of $f_{\varepsilon,X}(0,x)$. Under the homoscedasticity restriction, Theorem 2 implies that the standardized version $\widehat{\Gamma}(\tau):=Th^{d/2}\hat{H}^w(\tau)/\widehat{\sigma}_{0\tau}$ is asymptotically normal $N(0,1)$. This result forms the basis for the following one-sided asymptotic test for testing the null hypothesis of homoscedasticity at a given τ-quantile. For a given significance level α, we reject the null if $\widehat{\Gamma}(\tau) > z_\alpha$, where z_α is the one-sided critical value, i.e., the upper α-th percentile from the standard normal distribution. It is also worth noting that the result in Theorem 2 can be used to construct a valid confidence interval for $H(\tau)$ without the need to check the null of homoskedasticity, whereas the result in Theorem 1 can only be used when $H(\tau) \neq 0$.

Finally, the consistency and the sensitivity analysis of the test $\widehat{\Gamma}(\tau)$ to certain types of local alternatives are omitted, but the results for the case of the non-constant mean function $m(\cdot)$ are discussed in the next subsection and can be found in Proposition 1 and Theorem 5, respectively.

3.2 Non-constant mean

We now provide an estimator for the measure of heteroskedasticity at a given fixed quantile in the presence of non-constant and unknown mean $m(X_t)$. In particular, we use the sample $\{(Y_t, X_t)\}_{t=1}^{T}$ to nonparametrically estimate the following general measure of heteroskedasticity:

$$H^w(m,\tau) = 1 - \frac{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\phi}(X_t,\tau)\right) w(X_t)\right]}{E\left[\rho_\tau\left(\overline{Y}_t - \overline{\xi}(\tau)\right) w(X_t)\right]}, \text{ for } \tau \in (0,1),$$

where $\overline{Y}_t = Y_t - m(X_t)$. In practice, however, the functional form of $m(X_t)$ is unknown, but it can be estimated using the Nadaraya (1964) and Watson (1964) estimator:

$$\widehat{m}^{NW}(x) = \frac{\frac{1}{Tb^d}\sum_{t=1}^{T} K\left(\frac{x-X_t}{b}\right) Y_t}{\frac{1}{Tb^d}\sum_{t=1}^{T} K\left(\frac{x-X_t}{b}\right)}, \tag{13}$$

where $b \equiv b_T \in \mathbb{R}^+$ is a bandwidth parameter shrinking to 0 at a suitable rate as T diverges to infinity. We can also replace the Nadaraya-Watson estimator with a local polynomial type estimator $\widehat{m}^{LP}(x)$ as the one reported in Equation (11). It is worthwhile to mention that the bandwidth b needs to have a slower than h rate to annihilate the pre-estimation effect arising from estimating regression function $m(X_t)$ in the first stage. The detailed assumptions on both bandwidth parameters can be found in Appendix A.2.1.

Once we obtain $\widehat{m}(X_t)$, we then consider the feasible transformation of the dependent variable Y_t that will allow us to follow exactly the aforementioned procedure in Subsection 3.1. In other words, we use the following feasible (estimated) dependent variable $\widehat{\overline{Y}}_t := Y_t - \widehat{m}(X_t)$, where $\widehat{m}(X_t)$ can either be the Nadaraya-Watson estimator $\widehat{m}^{NW}(X_t)$ or the local polynomial estimator $\widehat{m}^{LP}(X_t)$. Next, the nonparametric estimator of our measure, say $\hat{H}^w(\widehat{m},\tau)$, is easily constructed through the following three steps:

(i) Given the sample $\{(Y_t, X_t)\}_{t=1}^{T}$ and a nonparametric estimator of $m(X_t)$, say $\widehat{m}^{NW}(X_t)$, we first generate the filtered sample $\left\{\left(\widehat{\overline{Y}}_t, X_t\right)\right\}_{t=1}^{T}$ for $\widehat{\overline{Y}}_t = Y_t - \widehat{m}^{NW}(X_t)$. We then estimate the τ-th marginal quantile of $\widehat{\overline{Y}}_t$, which we denote by $\widehat{\overline{\xi}}(\tau)$, by minimizing the empirical check loss function $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\overline{Y}}_t - \overline{\theta}\right)$ with respect to θ and compute the minimized value $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\overline{Y}}_t - \widehat{\overline{\xi}}(\tau)\right) w(X_t)$ for a given weight $w(\cdot)$.

(ii) Estimate the τ-th quantile regression of $\widehat{\overline{Y}}_t$ on X_t by minimizing the following objective function:

$$\frac{1}{T}\sum_{t=1}^{T} K_h\left(X_t - x\right)\rho_\tau\left(\widehat{\overline{Y}}_t - \sum_{0\le|\underline{r}|\le p}\overline{\beta}_{\underline{r}}(X_t - x)^{\underline{r}}\right). \tag{14}$$

Like in (11), denote the local polynomial estimator of the unconstrained quantile regression function $\overline{\phi}\left(x,\tau\right)$ by $\widehat{\overline{\phi}}(x,\tau) = \widehat{\overline{\beta}}_0(x)$. Then compute the minimized value by $T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\overline{Y}}_t - \widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)$, where $\widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)$ denotes the leave-observation-t-out estimator for $\overline{\phi}\left(X_t,\tau\right)$.

(iii) Finally, estimate the measure

$$\hat{H}^w\left(\widehat{m},\tau\right) = 1 - \frac{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\overline{Y}}_t - \widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)}{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(\widehat{\overline{Y}}_t - \widehat{\overline{\xi}}\left(\tau\right)\right)w\left(X_t\right)}. \tag{15}$$

We now state the main results when the mean function $m(X_t)$ is not constant. The following Theorem 3 provides an asymptotic Bahadur representation for the estimator $\hat{H}^w\left(\widehat{m},\tau\right)$ in (15); see the proof of Theorem 3 in Appendix A.2.1.

Theorem 3 *Assume that the mean function $m\left(X_t\right)$ in the regression* (1) *is not constant and unknown. Suppose Assumptions* ***C.1-C.10*** *in Appendix A.2.1 hold. If, furthermore, $p > d/2-1$, $h = O(T^{-\kappa})$, with $1/(2p+2+d) < \kappa < 1/(2d)$, then for a given $\tau \in \mathcal{T}$, we have*

$$\sqrt{T}\left(\hat{H}^w\left(\widehat{m},\tau\right) - H^w\left(m,\tau\right)\right) = \left(1 - H^w\left(m,\tau\right)\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\overline{e}_t - \overline{u}_t\right) + o_p(1),$$

where

$$\begin{aligned}
\overline{e}_t &= \frac{\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\xi}\left(\tau\right)\right)w\left(X_t\right) - E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\xi}\left(\tau\right)\right)w\left(X_t\right)\right]}{E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\xi}\left(\tau\right)\right)w\left(X_t\right)\right]},\\
\overline{u}_t &= \frac{\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right) - E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right]}{E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right]}.
\end{aligned}$$

Using the Bahadur representation in Theorem 3, we can immediately see that the nonparametric estimator $\hat{H}^w\left(\widehat{m},\tau\right)$ in (15) is consistent. Furthermore, applying the central limit theorem on the weakly dependent process $\left\{\left(\overline{e}_t - \overline{u}_t\right)\right\}$, Theorem 3 can be used to construct confidence intervals for the measure $H^w\left(m,\tau\right)$. However, when the homoscedasticity holds ($H^w\left(m,\tau\right) = 0$), we see from the Bahadur representation that the asymptotic variance of $\hat{H}^w\left(\widehat{m},\tau\right)$ degenerates to 0. Thus, the limiting distribution of $\hat{H}^w\left(\widehat{m},\tau\right)$ degenerates under the null of homoscedasticity. Using the standard theory for U-statistics, the next theorem provides the limiting distribution of $\hat{H}^w\left(\widehat{m},\tau\right)$ when homoscedasticity holds; see the proof of Theorem 4 in Appendix A.2.1.

Theorem 4 *Assume that the mean function $m(X_t)$ in the regression (1) is not constant and unknown. Suppose Assumptions **C.1-C.10** in Appendix A.2.1 hold. If, furthermore, $p > d/2 - 1$, $h = O(T^{-\kappa})$, with $1/(2p+2+d) < \kappa < 1/(2d)$, then under the null hypothesis of homoscedasticity, for a given $\tau \in \mathcal{T}$, we have*

$$Th^{d/2}\hat{H}^w(\widehat{m},\tau) \to_d N(0,\overline{\sigma}_{0\tau}^2),$$

with

$$\overline{\sigma}_{0\tau}^2 = \frac{2\tau^2(1-\tau)^2}{\overline{r}_\tau^2}\int K^2(u)\,du \int \frac{w^2(x)}{f_{\overline{\varepsilon},X}^2(0,x)} f_X^2(x)\,dx,$$

where $\overline{r}_\tau = E\left[\rho_\tau\left(Y_t - m(X_t) - \overline{\phi}(X_t,\tau)\right)w(X_t)\right]$, $f_{\overline{\varepsilon},X}(0,x)$ is the joint density of quantile error $\overline{\varepsilon}_t = Y_t - m(X_t) - \overline{\phi}(X_t,\tau)$ and X_t evaluated at $\overline{\varepsilon}_t = 0$, and $f_X(x)$ is the marginal density of X_t.

Likewise, a consistent estimator of $\overline{\sigma}_{0\tau}^2$ is given by

$$\widehat{\overline{\sigma}}_{0\tau}^2 = \frac{2\tau^2(1-\tau)^2}{\widehat{\overline{r}}_\tau^2}\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\frac{w^2(X_t)}{\widehat{f}_{\widehat{\overline{\varepsilon}},X}^2(0,X_t)}\frac{1}{h^d}K^2\left(\frac{X_t-X_s}{h}\right),$$

where

$$\widehat{\overline{r}}_\tau = \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\phi}}_{-t}(X_t,\tau)\right)w(X_t)$$

is a consistent estimator of $\overline{r}_\tau = E\left[\rho_\tau\left(Y_t - m(X_t) - \overline{\phi}(X_t,\tau)\right)w(X_t)\right]$, and $\widehat{f}_{\widehat{\overline{\varepsilon}},X}(0,x)$ is the kernel density estimator of $f_{\overline{\varepsilon},X}(0,x)$. It is worthwhile to remark that if we choose $w(x) = f_{\overline{\varepsilon},X}(0,x)$, our expressions for $\overline{\sigma}_{0\tau}^2$ and $\widehat{\overline{\sigma}}_{0\tau}^2$ reduce respectively to those for σ_0^2 and $\widehat{\sigma}_0^2$ in Jeong et al. (2012, Theorem 3.1(i) and (ii)) apart from the normalizing constants $\overline{r}_\tau^2$ and $\widehat{\overline{r}}_\tau^2$[2]. The consistency of $\widehat{\overline{\sigma}}_{0\tau}^2$ is shown in Appendix A.2.1. Furthermore, it is worthwhile to note that under Assumptions **C.1-C.10**, the nonparametric estimation of the mean function $m(X_t)$ is not affecting the limiting distribution of $\hat{H}^w(\widehat{m},\tau)$. In particular, under the assumption that the bandwidth parameter b in (13) converges to zero at a slower rate, the mean function $m(\cdot)$ can be treated as if it was known. Similar observations can also be found in e.g. Chan and Zhang (2011).

Under the homoscedasticity restriction, Theorem 4 implies that the standardized version $\widehat{\overline{\Gamma}}(\widehat{m},\tau) := Th^{d/2}\hat{H}^w(\widehat{m},\tau)/\widehat{\overline{\sigma}}_{0\tau}$ is asymptotically normal $N(0,1)$. This result forms the basis for the following one-sided asymptotic test for testing the null hypothesis of homoscedasticity at a given τ-quantile. For a given significance level α, we reject the null if $\widehat{\overline{\Gamma}}(\widehat{m},\tau) > z_\alpha$, where z_α is the one-sided critical value, i.e., the upper α-th percentile from the standard normal distribution. Recall that many other tests have been proposed in the literature to test if the variance function $\sigma(X_t)$ in Equation (1) is constant; see White (1980), Godfrey (1978), Machado and Santos Silva (2000), among others. In contrast, for the parametric quantile regression-based tests of heteroskedasticity, Koenker and Bassett (1982) and Newey and Powell

[2]Similar remarks hold for $\sigma_{0\tau}^2$ and $\widehat{\sigma}_{0\tau}^2$ in Theorem 2.

(1987) among others have proposed tests based on comparing different quantiles or expectile estimates. On the other hand, to name only a few, nonparametric regression-based tests of heteroskedasticity include the following: (1) Dette and Munk (1998), who propose a test based on an L_2 distance between the conditional variance and the constant variance; (2) Liero (2003), who provides a test based on an L_2 distance between the two nonparametric estimates of variance under heteroskedasticity and homoscedasticity; and (3) Zhu et al. (2001), who build a test of heteroskedasticity based on the integrated difference between the conditional variance and the unconditional variance weighted by an indicator function of covariate.

We next study the consistency and the power of the test $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ against global and local alternatives. The following proposition states the consistency of the test under a global alternative (see the proof of Proposition 1 in Appendix A.2.1):

Proposition 1 *Assume that the mean function $m(X_t)$ in the regression* (1) *is not constant and unknown. Suppose Assumptions **C.1-C.10** in Appendix A.2.1 hold. If, furthermore, $p > d/2-1$ and $h = O(T^{-\kappa})$, with $1/(2p+2+d) < \kappa < 1/(2d)$, then under the alternative hypothesis of heteroskedasticity [$H_A : \overline{Q}^{(\tau)}\left(\overline{Y}_t \middle| X_t\right) = \overline{\phi}(X_t,\tau)$], for a given $\tau \in \mathcal{T}$, we have*

$$Pr\left\{Th^{d/2}\hat{H}^w(\widehat{m},\tau)/\widehat{\overline{\sigma}}_{0\tau} > B_T\right\} \to 1,$$

for any non-stochastic sequence $B_T = o\left(Th^{d/2}\right)$.

Proposition 1 indicates that $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ diverges to positive infinity under heteroskedasticity and, therefore, is consistent against all global alternatives. We now examine the power of this test against Pitman local alternatives that approach the null at a proper rate. Specifically, we consider the local alternatives:

$$H_{1T} : \sigma(x) = \sigma_0 + \frac{1}{T^{1/2}h^{d/4}}\Delta_T(x) \text{ a.e.,} \tag{16}$$

for some non-negative and non-constant continuous function $\Delta_T(\cdot)$ for every $T \geq 1$. Notice that the formulation of local alternatives in (16) depends on σ instead of σ^2, which is slightly different from what is considered in Hsiao and Li (2001) and Su and Ullah (2013). This formulation greatly facilitates our local power analysis. The following Theorem 5 establishes the asymptotic local power property of the test $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ under the local alternatives in (16); see the proof of Theorem 5 in Appendix A.2.1.

Theorem 5 *Assume that the mean function $m(X_t)$ in the regression* (1) *is not constant and unknown. Suppose Assumptions **C.1-C.10** in Appendix A.2.1 hold. If, furthermore, $p > d/2-1$, $h = O(T^{-\kappa})$ with $1/(2p+2+d) < \kappa < 1/(2d)$, then under the local alternatives H_{1T} in (16), for a given $\tau \in \mathcal{T}$, we have*

$$Th^{d/2}\hat{H}^w(\widehat{m},\tau) \xrightarrow{d} N(\gamma, \overline{\sigma}^2_{0\tau}),$$

where $\overline{r}_\tau$ and $\overline{\sigma}^2_{0\tau}$ are defined in Theorem 4, and

$$\gamma = \overline{r}_\tau^{-1} \left(D^{-1}(\tau)\right)^2 \lim_{T\to\infty} E\left[\Delta_T^2(X_t)\, w(X_t)\, f_{\overline{\varepsilon}|X}(0|X_t)\right] > 0,$$

with $D^{-1}(\tau)$ as the τ-th quantile of the error term ϵ_t in the regression (1).

Theorem 5 shows that the limiting distribution of the estimator $\hat{H}^w(\widehat{m},\tau)$ is non-trivially shifted as $\gamma > 0$; therefore, the test $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ is able to detect Pitman local alternatives that converge to the null at the typical rate of $O_p\left(T^{-1/2}h^{-d/4}\right)$. In particular, the local power of the test increases with the deviation of γ.

4 Bootstrap

The results in Theorems 2 and 4 are valid only asymptotically, and the asymptotic normal distribution might not work well in finite samples. Particularly, for high-dimensional random variables, the asymptotic test is subject to size distortion because of possible finite sample bias in the nonparametric estimation due to the curse of dimensionality. Consequently, though it is asymptotically pivotal, the test $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ is severely distorted in finite samples when using standard normal critical values and is typically sensitive to the choice of the bandwidth. To overcome these problems, we introduce a bootstrap-based procedure in this section that approximates well the finite sample distribution of the test statistic $\widehat{\overline{\Gamma}}(\widehat{m},\tau)$ under the null. Following Su and Ullah (2013), we use a fixed regressor bootstrap method in the spirit of Hansen (2000), which does not aim to reproduce the whole dependence structure of the stochastic processes that generate the original sample but only a particular feature of it. The fixed regressor bootstrap is implemented as follows:

(1) For a given sample $\{(Y_t, X_t)\}_{t=1}^T$, perform a nonparametric regression of Y_t on X_t and obtain the nonparametric residuals $\widehat{\overline{Y}}_t = Y_t - \widehat{m}(X_t)$ for $t = 1, \ldots, T$. Then, compute the test statistic $\widehat{\overline{\Gamma}}(\widehat{m},\tau) = Th^{d/2}\hat{H}^w(\widehat{m},\tau)/\widehat{\overline{\sigma}}_{0\tau}$, for a given $\tau \in (0,1)$;

(2) For $t = 1, \ldots, T$, obtain the bootstrapped errors $\overline{Y}_t^*$ by random sampling with replacement from $\left\{\widehat{\overline{Y}}_s - \mu_{\overline{Y}},\ \text{for } s = 1, \ldots, T\right\}$, where $\mu_{\overline{Y}} = T^{-1}\sum_{s=1}^T \widehat{\overline{Y}}_s$ is the sample average of $\widehat{\overline{Y}}_s$. Then, generate the bootstrap analog of Y_t by holding X_t fixed: $Y_t^* = \widehat{m}(X_t) + \overline{Y}_t^*$, for $t = 1, \ldots, T$;

(3) Using the bootstrapped sample $\{(Y_t^*, X_t)\}_{t=1}^T$, perform a nonparametric regression of Y_t^* on X_t to obtain the bootstrapped nonparametric residuals $\widehat{\overline{Y}}_t^* = Y_t^* - \widehat{m}^*(X_t)$ for $t = 1, \ldots, T$. Then, compute the bootstrapped test statistic $\widehat{\overline{\Gamma}}^*(\widehat{m}^*,\tau) = Th^{d/2}\hat{H}^{w*}(\widehat{m}^*,\tau)/\widehat{\overline{\sigma}}_{0\tau}^*$, where $\hat{H}^{w*}(\widehat{m}^*,\tau)$ and $\widehat{\overline{\sigma}}_{0\tau}^*$ are defined analogously to $\hat{H}^w(\widehat{m},\tau)$ and $\widehat{\overline{\sigma}}_{0\tau}$ with $\widehat{\overline{Y}}_t$ being replaced by $\widehat{\overline{Y}}_t^*$;

(4) Repeat steps **(2)**-**(3)** B times so that we get a sample of bootstrapped statistics as $\left\{\widehat{\overline{\Gamma}}_j^*(\widehat{m_j}^*,\tau)\right\}_{j=1}^B$;

(5) Compute the bootstrapped p-value using $p^* = B^{-1}\sum_{j=1}^B 1\left(\widehat{\overline{\Gamma}}_j^*(\widehat{m_j}^*,\tau) > \widehat{\overline{\Gamma}}(\widehat{m},\tau)\right)$, and for a given significance level α, reject the null hypothesis if $p^* < \alpha$.

The following Theorem 6 establishes the asymptotic validity of the above fixed regressor bootstrap-based procedure (see the proof of Theorem 6 in Appendix A.2.1):

Theorem 6 *Suppose that the assumptions in Theorem 4 hold. Then, for a given $\tau \in \mathcal{T}$, we have*

$$\widehat{\widehat{\Gamma}}^{*}(\widehat{m}^{*},\tau) := \frac{Th^{d/2}\hat{H}^{w*}(\widehat{m}^{*},\tau)}{\widehat{\widehat{\sigma}}^{*}_{0\tau}} \xrightarrow{d} N(0,1)$$

conditionally on $\{(Y_t, X_t)\}_{t=1}^{T}$ as $T \to \infty$, where $\widehat{\widehat{\sigma}}^{*2}_{0\tau}$ is analogously defined as in Theorem 4.

The result in Theorem 6 provides an asymptotically valid approximation to the limiting null distribution of $\widehat{\widehat{\Gamma}}(\widehat{m},\tau)$. This result holds regardless of whether the null hypothesis is true or not. In the next section, we use Monte Carlo simulations to examine the performance of the bootstrap-based test in Theorem 6 for small to moderate-sized samples. We also examine the performance of another type of bootstrap, which is the smoothed local bootstrap; see Appendix A.2.1.

5 Monte Carlo simulations

We conduct a Monte Carlo simulation study to investigate the performance of the bootstrap-based test in Theorem 6 for testing the statistical significance of the measure of heteroskedasticity at a given fixed quantile. Since the non-constant mean case is the most relevant case in practice, in our simulations, we focus on testing the null hypothesis $H_0 : H^w(m,\tau) = 0$.

Though the asymptotic-based test $\widehat{\widehat{\Gamma}}(\widehat{m},\tau) = Th^{d/2}\hat{H}^w(\widehat{m},\tau)/\widehat{\widehat{\sigma}}_{0\tau}$, whose distribution is reported in Theorem 4, is not time consuming and is easy to implement, in small samples, its empirical size may differ significantly from the significance level. However, it is expected that the fixed regressor bootstrap-based test will help eliminate or mitigate the asymptotically negligible higher order terms that may have substantial adverse effects on the size of $\widehat{\widehat{\Gamma}}(\widehat{m},\tau)$. Thus, the objective is to assess the empirical size and power of the bootstrap-based test using several data generating processes (DGPs) that we present in Table 1, which we take from Su and Ullah (2013).

The six DGPs in Table 1 will be used to evaluate the empirical size and power of the bootstrap-based test in Theorem 6. The first two DGPs [DGP S1 and DGP S2] are used to investigate the size property since in these DGPs the null hypothesis (homoscedasticity) is satisfied. However, from DGP P1 to DGP P4, the null is not satisfied; therefore, they serve to illustrate the power of our test. Some of these DGPs [DGP S1, DGP P1, and DGP P2] can be viewed as if they represent cross-sectional data and others [DGP S2, DGP P3, and DGP P4] correspond to time series data. The time series processes are strictly stationary and ergodic. Furthermore, in DGP P4 we have $U_t = Y_t - 0.5Y_{t-1}$, which indicates that the dependent variable Y_t follows an AR-ARCH process. Following Su and Ullah (2013), in DGPs P1 to P4, we set $\delta_T = T^{-9/20}$ to study the local power behavior of our test under the Pitman local alternatives H_{1T} stated in (16).

Table 1: Data generating processes

DGPs	Variables of Interest		Conditional Variance of Y
	Y_t	X_t	$\sigma(X_t)$
DGP S1	$Y_t = 1 + X_t + \epsilon_t$	i.i.d. $U(-\sqrt{3}, \sqrt{3})$	1
DGP S2	$Y_t = 0.5Y_{t-1} + \epsilon_t$	Y_{t-1}	1
DGP P1	$Y_t = 1 + X_t + \sigma(X_t)\epsilon_t$	i.i.d. $U(-\sqrt{3}, \sqrt{3})$	$0.5 + \delta_T(X_t - 1)^2$
DGP P2	$Y_t = 1 + X_t + \sigma(X_t)\epsilon_t$	i.i.d. $U(-\sqrt{3}, \sqrt{3})$	$0.2 + \delta_T \exp(X_t)$
DGP P3	$Y_t = 0.5Y_{t-1} + \sigma(X_t)\epsilon_t$	Y_{t-1}	$0.1 + 5\exp(-\delta_T Y_{t-1}^2)$
DGP P4	$Y_t = 0.5Y_{t-1} + \sigma(X_t)\epsilon_t$	Y_{t-1}	$0.1 + 4\delta_T U_{t-1}^2$, where $U_t = Y_t - 0.5Y_{t-1}$

Note: This table summarizes the DGPs that we consider in the simulation study to investigate the properties (size and power) of the nonparametric test of measure of heteroskedasticity at a given fixed quantile. We simulate Y_t and X_t, for $t = 1, \ldots, T$, under the assumption that ϵ_t are i.i.d. $N(0, 1)$ [we also consider the case where ϵ_t are i.i.d. t_3] and $X_t \perp \epsilon_t$. The last column of the table reports the conditional variance of Y. When the latter is constant, then we are in the presence of homoscedasticity; when it is not, this corresponds to heteroskedasticity.

Table 2: Empirical rejection frequency of fixed regressor bootstrap when $\epsilon \sim N(0,1)$

Quantiles	DGPs					
	DGP S1	DGP S2	DGP P1	DGP P2	DGP P3	DGP P4
			$T = 50$			
$\tau = 0.25$	6.2	4.2	40.4	51.6	56.5	43.6
$\tau = 0.75$	4.8	3.4	55.2	66.4	70.0	56.7
			$T = 100$			
$\tau = 0.25$	5.8	3.2	55.1	69.5	62.4	76.5
$\tau = 0.75$	5.8	3.2	60.2	75.4	70.2	83.4
			$T = 200$			
$\tau = 0.25$	5.0	3.0	71.6	87.2	75.7	80.4
$\tau = 0.75$	5.2	3.0	75.2	88.3	77.8	85.2

Note: This table reports the empirical size and power of the fixed regressor bootstrap-based test, $\widehat{\widehat{\Gamma}}^{*}(\widehat{m}^{*}, \tau)$, in Theorem 6 for testing that the measure of heteroskedasticity at a given fixed quantile is equal to 0 $[H_0 : H^w(m, \tau) = 0]$ at $\alpha = 5\%$ significance level. The number of simulations is equal to 500 and the number of bootstrap resamples is $B = 199$. The error terms ϵ_t in regression (1) are i.i.d. $N(0,1)$.

The weight function $w(\cdot)$ in the estimator of the measure in Equation (15) is set to be equal to 1 everywhere, i.e., the trivial weight function, since we expect that the test's performance will not depend on this weight function. Furthermore, to estimate the nonparametric mean regression and the restricted and unrestricted quantile regression functions, we take the univariate kernel function $K(\cdot)$ equal to the standard normal density. Thereafter, we choose the two bandwidth sequences b and h by a "rule of thumb". For the first stage nonparametric mean regression, the bandwidth parameter is given by the following: $b = 1.5c{*}std(X_t){*}T^{-1/4}$. For the second stage restricted and unrestricted quantile regressions, the bandwidth is given by the following: $h = c * std(X_t) * T^{-1/3}$, where $std(X_t)$ is the sample standard deviation of X_t. We have reported the simulation results for $c = 1$. The results for $c = 0.5$ and $c = 1.5$ are not reported (available upon request), but they are quantitatively similar to those obtained for $c = 1$.

Three sample sizes $T = 50$, 100, and 200 are considered and two different quantile levels are examined $\tau = 0.25$, 0.75. For each DGP, we first generate $T + 200$ observations and then discard the first 200 observations to minimize the potential adverse effects of the initial values. We use 500 simulations to compute the empirical size and power. For each simulation, we use $B = 199$ bootstrap replications to approximate the finite sample distribution of $\widehat{\widehat{\Gamma}}(\widehat{m}, \tau)$. Finally, we focus on the nominal size 5%.

Table 2 reports the empirical size and power of the test statistic $\widehat{\widehat{\Gamma}}^*(\widehat{m}^*, \tau)$ when the error terms ϵ_t are i.i.d. $N(0, 1)$. As expected, the fixed regressor bootstrap test controls its size well for both small and moderate samples. Regarding the power, the test has reasonable power compared with various alternatives, even when the sample size is equal to 50; it also increases with sample size.

To compare with the above results, we also consider the case of heavy-tailed innovations: $\epsilon_t \sim i.i.d.$ t_3, where t_3 is a Student's t-distribution with three degrees of freedom. The new simulation results are reported in Table 3. From this, we again see that the fixed regressor bootstrap test controls its size and has reasonable power. This test is fairly robust to heavy-tailed distributions, and thus, it would be more appropriate to apply it to detect and measure heteroskedasticity in financial variables (returns) that are known to be leptokurtic.

Lastly, we perform additional simulations where the fixed regressor bootstrap is replaced by the smoothed local bootstrap; see Paparoditis and Politis (2000). One major advantage of the smoothed local bootstrap procedure is that it can preserve the unknown dependence structure in the stochastic processes generating the original sample. The implementation of the smoothed local bootstrap-based test is described in Appendix A.2.1, and the simulation results are reported in Table 4. Examining the results in Tables 2 and 4, we see that the smoothed local bootstrap-based test does slightly better (in terms of size and power) than the fixed regressor bootstrap-based test.

6 Empirical applications

We provide two empirical applications where our measures are applied to quantify the degree of heteroskedasticity using real data on economic and financial variables. In the first application, we are interested in measuring the degree of heteroskedasticity of income conditional on the years of education, and in the second one, we quantify the degree of heteroskedasticity of 29 individual stocks and of the S&P 500 Index.

6.1 Application I: Heteroskedasticity of income conditional on the years of education

This first application aims to apply the measures introduced in the previous sections to quantify the degree of heteroskedasticity of income conditional on the years of education for U.S. male and female workers. In particular, we measure the degree of heteroskedasticity at a fixed quantile of income distribution conditional on different ranges of years of education. The data used was from the March 2009 Current Population Survey (CPS) conducted by the Bureau of Labor Statistics in the United States Department of Labor.

Table 3: Empirical rejection frequency of fixed regressor bootstrap when $\epsilon \sim t_3$

Quantiles	DGPs					
	DGP S1	DGP S2	DGP P1	DGP P2	DGP P3	DGP P4
			$T = 50$			
$\tau = 0.25$	6.1	4.8	50.4	51.3	54.5	46.6
$\tau = 0.75$	3.7	5.3	55.5	65.7	71.2	57.6
			$T = 100$			
$\tau = 0.25$	3.8	3.9	59.1	66.7	64.2	78.4
$\tau = 0.75$	5.5	3.6	62.2	73.2	69.7	83.4
			$T = 200$			
$\tau = 0.25$	5.3	4.8	71.5	85.3	73.5	82.5
$\tau = 0.75$	5.1	4.6	75.6	89.5	78.6	82.8

Note: This table reports the empirical size and power of the fixed regressor bootstrap-based test, $\widehat{\overline{\Gamma}}^{*}(\widehat{m}^{*}, \tau)$, in Theorem 6 for testing that the measure of heteroskedasticity at a given fixed quantile is equal to 0 $[H_0 : H(m, \tau) = 0]$ at $\alpha = 5\%$ significance level. The number of simulations is equal to 500 and the number of bootstrap resamples is $B = 199$. The error terms ϵ_t in regression (1) are i.i.d. t_3, where t_3 is a Student's t-distribution with three degrees of freedom.

Table 4: Empirical rejection frequency of local smoothed bootstrap when $\epsilon \sim N(0,1)$

Quantiles	DGPs					
	DGP S1	DGP S2	DGP P1	DGP P2	DGP P3	DGP P4
			$T = 50$			
$\tau = 0.25$	5.1	4.6	42.5	50.3	55.2	41.7
$\tau = 0.75$	5.5	3.9	53.4	65.6	73.4	57.3
			$T = 100$			
$\tau = 0.25$	4.8	4.2	56.6	67.9	64.5	75.8
$\tau = 0.75$	4.7	5.2	62.3	76.7	74.6	84.5
			$T = 200$			
$\tau = 0.25$	4.5	4.7	73.5	88.3	78.4	81.5
$\tau = 0.75$	5.1	4.8	75.5	89.2	77.5	84.8

Note: This table reports the empirical size and power of the smoothed local bootstrap-based test [see Appendix A.2.1] for testing that the measure of heteroskedasticity at a given fixed quantile is equal to 0 $[H_0 : H(m, \tau) = 0]$ at $\alpha = 5\%$ significance level. The number of simulations is equal to 500 and the number of bootstrap resamples is $B = 199$. The error terms ϵ_t in regression (1) are i.i.d. $N(0, 1)$.

The CPS provides data on labor force characteristics of the population, including the level of employment, unemployment, and earnings. The variables of interest are average hourly earnings (AHE) and the number of years of education (EDU). The sample comprises 2989 full-time U.S. workers (1658 males and 1331 females) aged between 29 and 30 years and having between 6 and 18 years of education as of 2008. We assume that the conditional mean of income is given by the following nonparametric regression:

$$ahe_i = m\left(edu_i\right) + \sigma\left(edu_i\right)\epsilon_t,$$

where ahe_i is the average hourly earning of an individual i and edu_i is his/her education level.

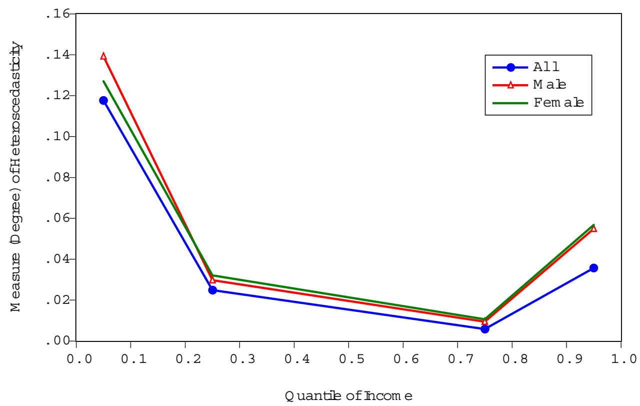
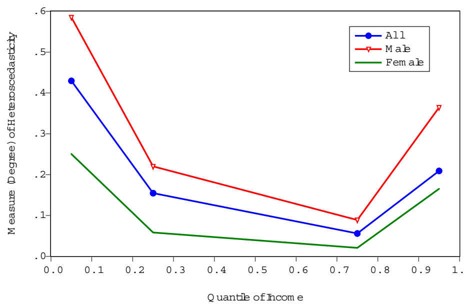
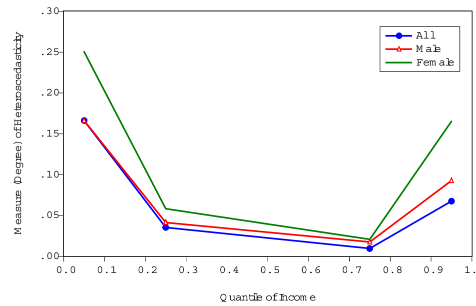


Figure 1: This figure provides the measure (degree) of heteroskedasticity at different quantiles of the conditional distribution of U.S. male and female income conditional on all the years of education, which corresponds to a weighting function $w\left(X_t\right) = 1$.

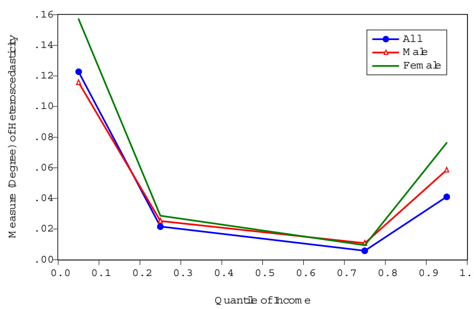
Figures 1 and 2 illustrate the results of estimating the measures of heteroskedasticity of income. We have applied the bootstrap-based test in Theorem 6 to test for the statistical significance of the estimates of the measures. The results were omitted as they indicate that all these estimates (at different quantiles of the income distribution and different ranges of the years of education) are statistically significant even at 1% significance level. On the one hand, the results in Figure 1 illustrate the case where the weighting function $w\left(X_t\right)$ is equal to 1. In other words, this shows the estimates of the degree of heteroskedasticity of income conditional on all years of education. On the other hand, the results in Figure 2 present the estimates of the measures of heteroskedasticity of income conditional on different ranges of the years of education.

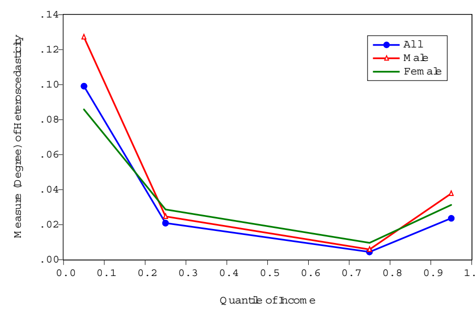
Conditional on all years of education, Figure 1 shows that it is difficult to distinguish between the degrees of measures of heteroskedasticity for males and females. However, when one considers different ranges of the years of education, the difference in the degree of income variation for males and females becomes clearer, especially for groups with low levels of education; see Figure 2. Furthermore, the degree of income variation



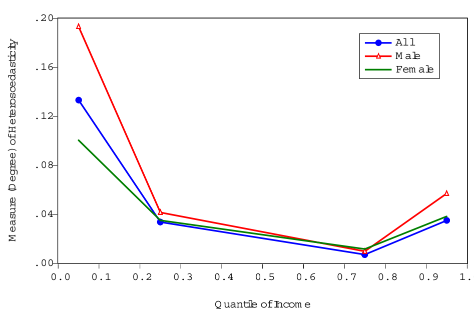


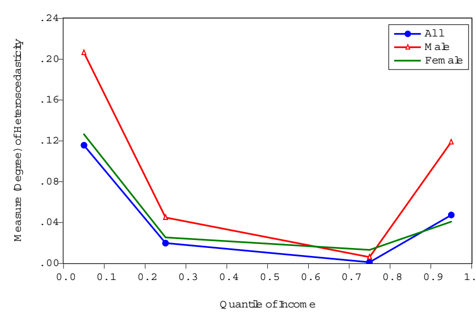
(a) Education level lower than the 5th quantile (b) Education level lower than its 25th quantile





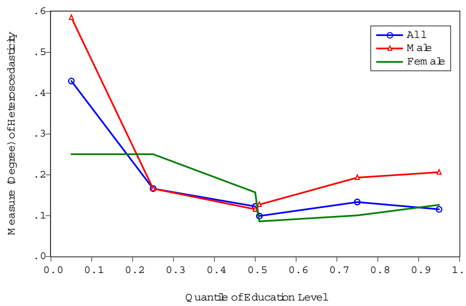
(c) Education level lower than its 50th quantile (d) Education level higher than its 50th quantile



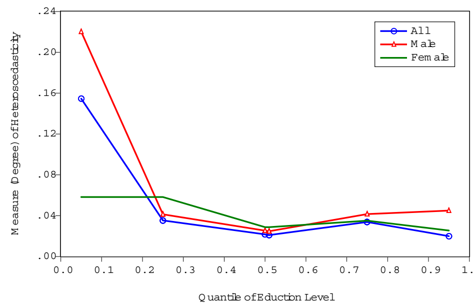


(e) Education level higher than its 75th quantile (f) Education level higher than its 95th quantile

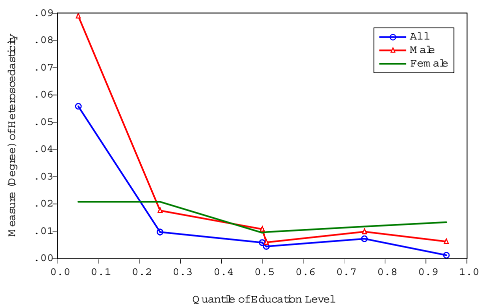
Figure 2: This figure provides the measure (degree) of heteroskedasticity at different quantiles of the conditional distribution of U.S. male and female income conditional on different quantile ranges of the distribution of years of education: (a) (0, 0.05), (b) (0, 0.25), (c) (0, 0.5), (d) (0.5, 1), (e) (0.75, 1), and (f) (0.95, 1). The quantile ranges define the weighting function $w(X_t)$. For example, for the quantile range (0, 0.05), $w(X_t) = \mathbb{I}[X_t \leq q^x(0.05)]$.



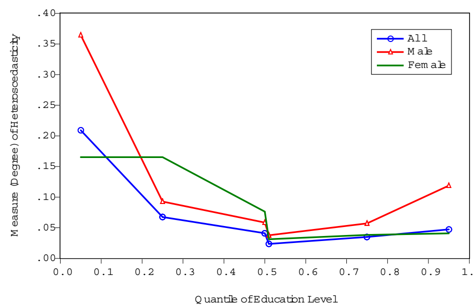
(a) 5th quantile of Income



(b) 25th quantile of Income



(c) 75th quantile of Income



(d) 95th quantile of Income

Figure 3: This figure provides the measure (degree) of heteroskedasticity at each quantile of U.S. male and female income (5%, 25%, 75%, and 95%) across different ranges of the distribution of years of education [in the figure "Quantile of Education Level"].

for females is generally smaller than the degree of income variation for males [see for example the sub-figure (a) of Figure 2].

For a given fixed quantile of income distribution (5%, 25%, 75%, or 95% quantiles), Figure 3 illustrates the degree of heteroskedasticity of income as a function of ranges of years of education. From this, we see that the degree of income variation decreases when the years of education increase, which is true for all quantiles under consideration. Thus, the income of highly educated people varies less compared to the income of low educated people. This is more apparent for male workers compared with female workers.

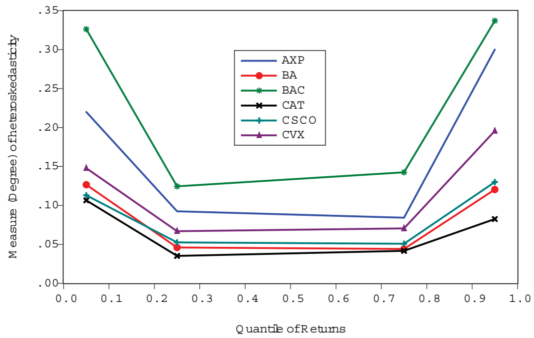
6.2 Application II: Heteroskedasticity of 30 stock returns

In this second application, we use the proposed measures to quantify the degree of heteroskedasticity of many stock returns. The dataset comes from Yahoo Finance and consists of 29 daily individual stocks and daily S&P 500 Index with 2517 observations over the period from January 1, 2007 to December 31, 2016. The 29 individual stocks are as follows: American Express Company (AXP); Boeing Company (BA); Bank of America Corporation (BAC); Caterpillar Inc. (CAT); Cisco Systems, Inc. (CSCO); Chevron Corporation (CVX); E. I. du Pont de Nemours and Company (DD); Walt Disney Company (DIS); General Electric Company (GE); Home Depot, Inc. (HD); Hewlett-Packard Company (HPQ); International Business Machines Corporation (IBM); Intel Corporation (INTC); Johnson & Johnson (JNJ); JPMorgan Chase & Co. (JPM); Coca-Cola Company (KO); McDonald's Corp. (MCD); 3M Company (MMM); Merck & Co. Inc. (MRK); Microsoft Corporation (MSFT); Pfizer Inc. (PFE); Procter & Gamble Co. (PG); AT&T, Inc. (T); Travelers Companies, Inc. (TRV); UnitedHealth Group Incorporated (UNH); United Technologies Corp. (UTX); Verizon Communications Inc. (VZ); Wal-Mart Stores Inc. (WMT); and Exxon Mobil Corporation (XOM). For each stock, we compute the continuously compounded daily returns, say r_t, by taking the difference between the logarithm of the price at time t and the logarithm of the price at time $t-1$. We assume that the conditional mean of each stock return is given by the following nonparametric regression:

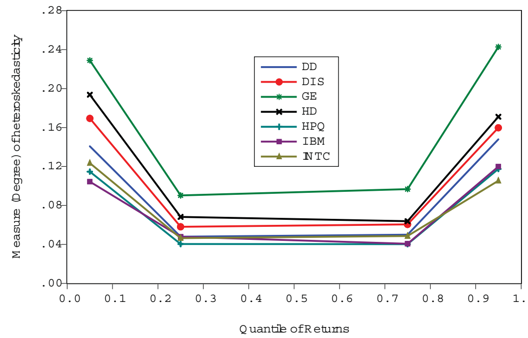
$$r_t = m\left(r_{t-1}\right) + \sigma\left(r_{t-1}\right)\epsilon_t,$$

where the conditional information is given by the past return r_{t-1}.

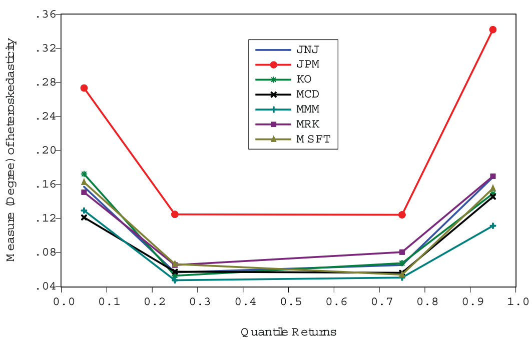
Figure 4 illustrates the results of estimating the measure of heteroskedasticity for each of the 30 stock returns and for different quantiles of their conditional distributions. We use the bootstrap-based test in Theorem 6 to test for the statistical significance of the estimates of the measures. We find (results are not reported but available upon request) that most of the bootstrapped p-values are close to zero, which confirms that stock returns are conditionally heteroskedastic. In addition, Figure 4 shows that there is some heterogeneity in the degree of heteroskedasticity across the 30 stocks under consideration.



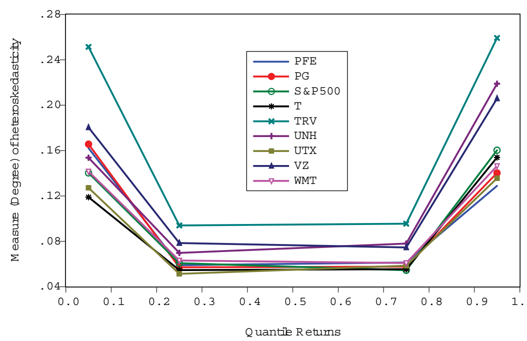
(a) Stocks: AXP, BA, BAC, CAT, CSCO, and CVX



(b) Stocks: DD, DIS, GE, HD, HPQ, IBM, and INTC



(c) Stocks: JNJ, JPM, KO, MCD, MMM, MRK, and MSFT



(d) Stocks: PFE, PG, SP500, T, TRV, UNH, UTX, VZ, and WMT

Figure 4: This figure provides the measure (degree) of heteroskedasticity at different quantiles of the conditional distributions of 30 daily stock returns (including the S&P 500 Index returns) conditional on their past returns. The weighting function $w(X_t) = 1$.

7 Conclusion

We introduced a measure to quantify the strength of heteroskedasticity at a given fixed quantile of the conditional distribution of a random variable conditional on the support or sub-support of other random variables. Our measure of heteroskedasticity is based on nonparametric quantile regressions and is expressed in terms of unrestricted and restricted expectations of quantile loss functions. It can be consistently estimated by replacing the unknown expectations by their nonparametric estimates. We derived a Bahadur-type representation for the nonparametric estimator of the measure. We provided the asymptotic distribution of this estimator, which one can use to build tests for the statistical significance of the measure. Thereafter, we established the validity of a fixed regressor bootstrap that one can use in finite-sample settings to perform tests. A Monte Carlo simulation study revealed that the bootstrap-based test has a good finite sample size and power for a variety of data generating processes and different sample sizes. Finally, two empirical applications were provided to illustrate the importance of the proposed measure. In the first application, we were interested in measuring the degree of heteroskedasticity of income conditional on the years of education, and in the second one, we quantified the degree of heteroskedasticity of 30 stock returns. For the first application, our results showed that the degree of income variation decreases when the years of education increase. Thus, the income of highly educated people varies less compared with the income of those with low levels of education. Furthermore, we found that the degree of income variation for females is generally smaller than the degree of income variation for males. For the second application, the results confirmed that all stock returns under consideration are conditionally heteroskedastic. In addition, these results showed that there is some heterogeneity in the degree of heteroskedasticity across the stocks. Finally, for all 30 stocks we used, we found that the degree of heteroskedasticity is high at the extremes of the conditional distribution of returns.

References

[1] Arabmazar, A. and Schmidt, P. (1981). "Further evidence on the robustness of the Tobit estimator to heteroskedasticity," *Journal of Econometrics*, vol. 17(2), pp. 253–258.

[2] Cattaneo, M.D., Jansson, M. and Newey, W.K. (2018). "Inference in linear regression models with many covariates and heteroskedasticity," *Journal of the American Statistical Association*, vol. 113(523), 1350–1361.

[3] Chan, N.H. and Zhang, R.-M. (2011). "Quantile inference for heteroscedastic regression models," *Journal of Statistical Planning and Inference*, vol. 141, pp. 2079–2090

[4] Chaudhuri, P. (1991). "Nonparametric estimates of regression quantiles and their local Bahadur representation," *The Annals of Statistics*, vol. 19, pp. 760–777.

[5] Cribari-Neto, F. (2004). "Asymptotic inference under heteroskedasticity of unknown form," *Computational Statistics & Data Analysis*, vol. 45(2), pp. 215–233.

[6] Downs, G.W. and Rocke, D.M. (1979). "Interpreting heteroscedasticity," *American Journal of Political Science*, pp. 816–828.

[7] Dette, H. and Munk, A. (1998). "Testing heteroscedasticity in nonparametric regression," *Journal of the Royal Statistical Society, Series B*, vol. 60, pp. 693–708.

[8] Dufour, J.-M. (2003). "Identification, weak instruments and statistical inference in econometrics," *Canadian Journal of Economics*, vol. 36 (4), pp. 767–808.

[9] Dufour, J.M. and Taamouti, A. (2010). "Exact optimal inference in regression models under heteroskedasticity and non-normality of unknown form," *Computational Statistics & Data Analysis*, vol. 54(11), pp. 2532–2553.

[10] Evans, M.A. and King, M.L. (1988). "A further class of tests for heteroscedasticity,"*Journal of Econometrics*, vol. 37, pp. 265–276.

[11] Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications*, Chapman & Hall, London.

[12] Gao, J. (2007). *Nonlinear time series: semiparametric and nonparametric methods*, Chapman & Hall, CRC.

[13] Godfrey, L.G. (1978). "Testing for multiplicative heteroskedasticity," *Journal of Econometrics*, vol. 8(2), pp. 227–236.

[14] Guerre, E. and Sabbah, C. (2012). "Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function," *Econometric Theory*, vol. 28, pp. 87–129.

[15] Hall, P. (1984). "Central limit theorem for integrated square error of multivariate nonparametric density estimators," *Journal of Multivariate Analysis*, vol. 14, pp. 1–16.

[16] Hall, P., Peng, L. and Yao, Q. (2002). "Prediction and nonparametric estimation for time series with heavy tails," *Journal of Time Series Analysis*, vol. 23, pp. 313–331.

[17] Hansen, B.E. (2000). "Testing for structural change in conditional models," *Journal of Econometrics*, vol. 97, pp. 93–115.

[18] Hausman, J. and Palmer, C. (2012). "Heteroskedasticity–robust inference in finite samples," *Economics Letters*, vol. 116(2), pp. 232–235.

[19] Honda, T. (2000). "Nonparametric estimation of a conditional quantile for α-mixing processes," *The Annals of the Institute of Statistical Mathematics*, vol. 52, pp. 459–470.

[20] Hsiao, C. and Li, Q. (2001). "A consistent test for conditional heteroskedasticity in time-series regression models," *Econometric Theory*, vol. 17, pp. 188–221.

[21] Hurd, M. (1979). "Estimation in truncated samples when there is heteroscedasticity," *Journal of Econometrics*, vol. 11, pp. 247–258.

[22] Koenker, R. and Bassett, G.S. (1982). "Robust tests for heteroscedasticity based on regression quantiles," *Econometrica*, vol. 50, pp. 43–61.

[23] Koenker, R. (2005). *Quantile regression*, Cambridge University Press, Cambridge.

[24] Koenker, R. and Zhao, Q. (1996). "Conditional quantile estimation and inference for ARCH models," *Econometric Theory*, vol. 12, pp. 793–813.

[25] Koenker, R. and Machado, J. A. (1999). "Goodness of fit and related inference processes for quantile regression," *Journal of the American Statistical Association*, vol. 94, pp. 1296–1310.

[26] Kong, E., Linton, O., Xia, Y. (2010). "Uniform Bahadur representation for local polynomial estimates of M-regression and its application to the additive model," *Econometric Theory*, vol. 26, pp. 1529–1564.

[27] Lee, B. (1992). "A heteroskedasticity test robust to conditional mean misspecification," *Econometrica*, vol. 60, pp. 159–171.

[28] Liero, H. (2003). "Testing homoscedasticity in nonparametric regression," *Journal of Nonparametric Statistics*, vol. 15, pp. 31–51.

[29] Machado, J.A.F. and Santos Silva, J.M.C. (2000). "Glejser's test revisited," *Journal of Econometrics*, vol. 97, pp. 189–202.

[30] Meghir, C. and Pistaferri, L. (2004). "Income variance dynamics and heterogeneity," *Econometrica*, vol. 72(1), pp.1–32.

[31] Nadaraya, E.A. (1964). "On estimating regression," *Theory of Probability and its Applications*, vol. 9(1), pp. 141–142.

[32] Newey, W.K. and Powell, J.L. (1987). "Asymmetric least squares estimation and testing," *Econometrica*, pp. 819–847.

[33] Noh, H., El Ghouch, A., Van Keilegom, I. (2013). “Quality of fit measures in the framework of quantile regression,” *Scandinavian Journal of Statistics*, vol. 40(1), pp. 105–118.

[34] Paparoditis, E., and Politis, D. (2000). “The local bootstrap for kernel estimators under general dependence conditions,” *Annals of the Institute of Statistical Mathematics*, vol. 52, pp. 139–159.

[35] Ruppert, D. and Wand, M.P. (1994). “Multivariate locally weighted least squares regression,” *The Annals of Statistics*, vol. 22, pp. 1346–1370.

[36] Senyo, B. and Adjibolosoo, S.K. (1989). “Measuring the degree of severity of heteroskedasticity and the choice between the OLS estimator and the 2SAE,” *Communications in statistics-theory and methods*, vol. 18(9), pp. 3451–3462.

[37] Su, L. and Ullah, A. (2013). “A nonparametric goodness-of-fit-based test for conditional heteroskedasticity,” *Econometric Theory*, vol. 29, pp. 187–212.

[38] Sun, S. (2006). “The Bahadur representation for sample quantiles under weak dependence,” *Statistics and Probability Letters*, vol. 76, pp. 1238–1244.

[39] Watson, G.S. (1964). “Smooth regression analysis,” *Sankhya*, vol. 26, pp. 359–372.

[40] White, H. (1980). “A heteroskedasticity–consistent covariance matrix estimator and a direct test for heteroskedasticity,” *Econometrica*, pp. 817–838.

[41] Yu, K. and Jones, M. (1998). “Local linear quantile regression,” *Journal of the American Statistical Association*, vol. 93, pp. 228–237.

[42] Zheng, X. (2009). “Testing heteroscedasticity in nonlinear and nonparametric regressions,” *Canadian Journal of Statistics*, vol. 37, pp. 282–300.

[43] Zhu, L., Fujikoshi, Y., Naito, K. (2001). “Heteroscedasticity checks for regression models,” *Science in China Series A–Mathematics*, vol. 44, pp. 1236–1252.

A Appendix

In this appendix, we describe the implementation of smoothed local bootstrap test as an alternative to fixed-regressor bootstrap test. We also provide our assumptions and sketch proofs of the theoretical results.

A.1 Smoothed local bootstrap-based test

Here, we describe how a smoothed local bootstrap-based test for testing the null hypothesis $H_0 : H^w(m, \tau) = 0$ can be implemented. We first need to consider the following notations. In the sequel, $X \sim f_X$ means that the random variable X is generated from a density function f_X. Let L_1, L_2 and L_3 be three kernels (standard normal density) and h^* be a bandwidth parameter for the bootstrap. The local smoothed bootstrap is implemented in the following four steps:

(1) Draw a bootstrapped sample $\{(Y_t^*, X_t^*)\}_{t=1}^T$. We first draw X_{t-1}^* using its nonparametric density

$$X_{t-1}^* \sim \frac{1}{Th^{*d}} \sum_{s=1}^{T} L_1\left(\frac{X_{s-1} - x}{h^*}\right).$$

Then conditional on X_{t-1}^*, we draw X_t^*, and conditional on X_t^* we draw Y_t^* independently from the following two non-parametric conditional probability densities:

$$\begin{aligned} X_t^* &\sim \tfrac{1}{h^{*d}} \textstyle\sum_{s=1}^T L_1\left(\frac{X_{s-1}-X_{t-1}^*}{h^*}\right) L_2\left(\frac{X_s - x}{h^*}\right) / \textstyle\sum_{s=1}^T L_1\left(\frac{X_{s-1}-X_{t-1}^*}{h^*}\right), \\ Y_t^* &\sim \tfrac{1}{h^*} \textstyle\sum_{s=1}^T L_1\left(\frac{X_s - X_t^*}{h^*}\right) L_3\left(\frac{Y_s - y}{h^*}\right) / \textstyle\sum_{s=1}^T L_1\left(\frac{X_s - X_t^*}{h^*}\right); \end{aligned}$$

(2) Based on the bootstrapped sample $\{(Y_t^*, X_t^*)\}_{t=1}^T$, we compute the bootstrapped version of the test statistic: $\widehat{\overline{\Gamma}}_{LS}^{*}(\widehat{m}^*, \tau) = \frac{Th^{d/2}\hat{H}^*(\widehat{m}^*,\tau)}{\widehat{\overline{\sigma}}_{0\tau}^*}$;

(3) Repeat the steps **(1)**-**(2)** B times so that we get $\widehat{\overline{\Gamma}}_{LS,j}^{*}(\widehat{m}^*, \tau)$, for $j = 1, \ldots, B$;

(4) Compute the bootstrapped p-value using $p^* = B^{-1} \sum_{j=1}^B 1\left(\widehat{\overline{\Gamma}}_{LS,j}^{*}(\widehat{m}^*, \tau) > \widehat{\overline{\Gamma}}(\widehat{m}, \tau)\right)$, where $\widehat{\overline{\Gamma}}(\widehat{m}, \tau) = \frac{Th^{d/2}\hat{H}^w(\widehat{m},\tau)}{\widehat{\overline{\sigma}}_{0\tau}}$ is the test statistic based on the original sample $\{(Y_t, X_t)\}_{t=1}^T$, and for a given significance level α, we reject the null hypothesis if $p^* < \alpha$.

A.2 Assumptions and proofs of main results

In this appendix, we include our main assumptions and sketch proofs of the theoretical results. For the asymptotic properties, we only include the proofs for the non-constant mean case, as the proofs of Theorems 1 and 2 are similar but much simpler than those of Theorems 3 and 4.

A.2.1 Technical assumptions

Here, we provide the necessary assumptions needed to derive the theoretical results in the paper. These assumptions mainly deal with the time series data with certain dependence structure, but our results still valid for cross-sectional data. We consider a set of standard assumptions that have been widely used in the literature; see for example Kong et al. (2010) and Noh et al. (2013) among others.

Let $\{(X_t, Y_t)\}$ be a strongly mixing stationary process with $\gamma(k)$ its strong mixing coefficient satisfying:

$$\gamma(k) = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_k^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)| \to 0 \text{ as } k \to \infty,$$

with $\mathcal{F}_a^b = \sigma\left(\{(X_t, Y_t)\}_{t=a}^b\right)$, where $\sigma(\cdot)$ means the smallest sigma algebra. Furthermore, let V_x be an open convex set in $\mathbb{R}^d$. Define $\varphi(u, \tau) = \tau 1(u \geq 0) + (\tau - 1) 1(u < 0) = \tau - 1(u < 0)$ to be the piecewise constant derivative of the loss function $\rho_\tau(u)$, with $1(\cdot)$ the indicator function. We now consider the assumptions:

C.1. The processes $\{(X_t, Y_t)\}$ are strongly mixing with mixing coefficients $\gamma(k)$ satisfying

$$\sum_{k=1}^{\infty} k^\alpha \gamma(k)^{1-2/\nu}, \text{ for some } \nu > 2 \text{ and } \alpha > (p + d + 1)(1 - 2/\nu)/d.$$

C.2. All partial derivatives of $\overline{\phi}(x, \tau)$ up to order $p + 1$ exist and are continuous for all $x \in V_x$, and there exists a constant $C > 0$ such that $|D^{\underline{r}}\overline{\phi}(x, \tau)| \leq C$ for all $x \in V_x$ and $\underline{r} = p + 1$.

C.3. The marginal density of $\overline{\varepsilon}_t = Y_t - m(X_t) - \overline{\phi}(X_t, \tau)$ is bounded and $E(\varphi(\overline{\varepsilon}_t, \tau) | X_t) = 0$.

C.4. For all e in a neighbourhood of zero, the conditional density $f_{\overline{\varepsilon}|X}(e|x)$ of $\overline{\varepsilon}_t = Y_t - m(X_t) - \overline{\phi}(X_t, \tau)$ given $X_t = x$ satisfies

$$\left|f_{\overline{\varepsilon}|X}(e|x_1) - f_{\overline{\varepsilon}|X}(e|x_2)\right| \leq K_e \|x_1 - x_2\|,$$

where K_e is a positive constant depending on e. Further, the conditional density is positive for $e = 0$ for all values of $x \in V_x$, and its first partial derivative with respect to e, $D^1 f_{\overline{\varepsilon}|X}(e|x)$, is bounded for all $x \in V_x$ and e in a neighbourhood of zero.

C.5. The weight function $w(x)$ is continuous, and its support $\mathcal{D} \subset V_x$ is compact and has non-empty interior.

C.6. The kernel function $K(\cdot)$ has a compact support and $\left|H_{\underline{j}}(x_1) - H_{\underline{j}}(x_2)\right| \leq \|x_1 - x_2\|$ for all j with $0 \leq \underline{j} \leq 2p + 1$, where $H_{\underline{j}}(x) = x^{\underline{j}} K(x)$.

C.7. The marginal density function of X_t, $f_X(x)$, is positive and bounded with bounded first-order derivatives on V_x. The joint density of (X_1, X_{l+1}) satisfies $f(x_1, x_2; l) \leq C < \infty$ for all $l \geq 1$.

C.8. The conditional density $f_{X_0|X_1}$ of X_t given X_{t+1} exists and is bounded. The conditional density function $f_{(X_1, X_{l+1})|(X_2, X_l+2)}$ of (X_1, X_{l+1}) given (X_2, X_{l+2}) exists and is bounded for all $l \geq 1$.

C.9. The distribution function of Y_t, $F_Y(\cdot)$, has bounded second derivative in a neighbourhood of $m(X_t) + \overline{\phi}(X_t, \tau)$ and $f_Y\left(m(X_t) + \overline{\phi}(X_t, \tau)\right) > 0$, where $f_Y(\cdot)$ is the marginal density function of Y_t.

C.10. The bandwidth sequences h and b satisfy $h \to 0$, $Th^{d+2(p+1)}/\log T = O(1)$, $b \to 0$, and $Tb \to \infty$ as $T \to \infty$. Furthermore, we assume $Th^{2d}/(\log T)^3 \to \infty$, and $h = o(b)$.

A.2.2 Proofs

This appendix shows the proofs of the main theoretical results developed in Sections 3 and 4. The proofs of Theorems 1 and 2 are similar to but simpler than those of Theorems 3 and 4 and therefore they are omitted.

First of all, recall that when the function $m(\cdot)$ is unknown, the nonparametric estimator of our weighted measure of heteroskedasticity at a given fixed quantile τ is defined by

$$\hat{H}^w(\widehat{m},\tau) = 1 - \frac{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\phi}}_{-t}(X_t,\tau)\right)w(X_t)}{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\xi}}(\tau)\right)w(X_t)}, \text{ for } \tau \in (0,1),$$

where $\widehat{m}(X_t)$, $\widehat{\overline{\xi}}(\tau)$, and $\widehat{\overline{\phi}}_{-t}(X_t,\tau)$ are respectively the nonparametric estimators for the conditional mean function $m(X_t)$, the τ-th marginal quantile of $Y_t - m(X_t)$, and the τ-th conditional quantile function of $Y_t - m(X_t)$ given X_t leaving observation t out. In addition, let $\widehat{\overline{\phi}}(x,\tau)$ denote the estimator when evaluating on the general x. Four auxiliary lemmas which are useful to prove our main results are given below. The first two lemmas establish the Bahadur representations of $\widehat{\overline{\xi}}(\tau)$ and $\widehat{\overline{\phi}}(x,\tau)$, respectively.

Lemma 1: *Let $f_Y'(y)$ be bounded in a neighbourhood of $m(X_t)+\overline{\xi}(\tau)$. Then, with probability one, we have*

$$\widehat{\overline{\xi}}(\tau) - \overline{\xi}(\tau) = -\frac{F_{YT}\left(\overline{\xi}(\tau)\right) - \tau}{f_Y\left(\overline{\xi}(\tau)\right)} + R_T^*,$$

where $F_{YT}(y) = T^{-1}\sum_{t=1}^{T} 1(Y_t \le y)$ is the empirical distribution function and $R_T^* = o_p\left(T^{-3/4+\delta}\log T\right)$, for some $\delta \in (0,1/4)$.

Proof of Lemma 1: See the proofs of Theorems 1 and 2 of Sun (2006).■

Lemma 2: *Let e_1 be an $N\times 1$ vector with its first element given by 1 and all others 0. Suppose Assumptions **C.1-C.10** hold and $h = O(T^{-\kappa})$ with $\kappa > 1/(2q+2+d)$. Then, with probability one, we have*

$$\widehat{\overline{\phi}}(x,\tau) - \overline{\phi}(x,\tau) = -e_1'\frac{H_T^{-1}}{Th^d}S_{T,p}^{-1}(x)\sum_{t=1}^{T}K_h(X_t - x)\varphi(\overline{\varepsilon}_t)\mu(X_t - x) + R_T^*(x),$$

where $\overline{\varepsilon}_t = Y_t - m(X_t) - \overline{\phi}(X_t,\tau)$ is the unrestricted quantile error and $R_T^*(x) = o_p\left(\left(Th^d\right)^{-1/2}\right)$ uniformly in $x \in \mathcal{D}$ and $\mathcal{D}$ is the compact support of the weighting function $w(\cdot)$.

Proof of Lemma 2: This follows by the standard results for M-regression using local polynomial methods in Kong et al. (2010).■

The next two lemmas provide the Bahadur representations of the check loss functions $\rho_\tau(\cdot)$ involving $\widehat{\overline{\xi}}(\tau)$ and $\widehat{\overline{\phi}}(X_t,\tau)$, respectively. They are needed to prove Theorems 3 and 4.

Lemma 3: *Suppose Assumptions **C.1-C.10** hold. Then,*

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\xi}}(\tau)\right)w(X_t) - E\left[\rho_\tau\left(Y_t - m(X_t) - \overline{\xi}(\tau)\right)w(X_t)\right]\\ =&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(X_t - m(X_t) - \overline{\xi}(\tau)\right)w(X_t) - E\left[\rho_\tau\left(X_t - m(X_t) - \overline{\xi}(\tau)\right)w(X_t)\right] + o_p(T^{-1/2}).\end{aligned}$$

Proof of Lemma 3: The proof of Lemma 3 is omitted, because it can be regarded as a similar case of the following Lemma 4, with the estimated conditional quantile function $\widehat{\overline{\phi}}(X_t,\tau)$ replaced by the estimated marginal sample quantile $\widehat{\overline{\xi}}(\tau)$. Then, combining the Bahadur representation of $\widehat{\overline{\xi}}(\tau)$ in Lemma 1 and that of $\widehat{m}(X_t)$ will prove Lemma 3.■

Lemma 4: *Suppose Assumptions **C.1-C.10** hold, furthermore,* $p > d/2-1$ *and* $h = O(T^{-\kappa})$ *with* $1/(2p+2+d) < \kappa < 1/(2d)$*. Then,*

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t-\widehat{m}\left(X_t\right)-\widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)-E\left[\rho_\tau\left(Y_t-m\left(X_t\right)-\overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right]\\
=&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t-m\left(X_t\right)-\overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)-E\left[\rho_\tau\left(Y_t-m\left(X_t\right)-\overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right]+o_p(T^{-1/2}).
\end{aligned}$$

Proof of Lemma 4: Notice that for any x and y, we have

$$\rho_\tau\left(x-y\right)-\rho_\tau\left(x\right)=\left(-y\right)\varphi\left(x\right)+2\left(y-x\right)\left[1\left(y>x>0\right)-1\left(y<x<0\right)\right],$$

where $\varphi\left(x\right):=\varphi\left(x,\tau\right)=\tau-1\left(x<0\right)$ is the piecewise constant derivative of $\rho_\tau\left(x\right)$. Let $\widehat{d}\left(x\right)=\widehat{m}\left(x\right)-m\left(x\right)$ and $\widehat{\overline{d}}\left(x\right)=\widehat{\overline{\phi}}\left(x,\tau\right)-\overline{\phi}\left(x,\tau\right)$. Then,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t-\widehat{m}\left(X_t\right)-\widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)-\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t-m\left(X_t\right)-\overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\\
=&-\frac{1}{T}\sum_{t=1}^{T}\widehat{d}\left(X_t\right)w\left(X_t\right)\varphi\left(\overline{\varepsilon}_t\right)-\frac{1}{T}\sum_{t=1}^{T}\widehat{\overline{d}}\left(X_t\right)w\left(X_t\right)\varphi\left(\overline{\varepsilon}_t\right)\\
&-\frac{2}{T}\sum_{t=1}^{T}\left\{\overline{\varepsilon}_t-\left[\widehat{d}\left(X_t\right)+\widehat{\overline{d}}\left(X_t\right)\right]\right\}\left\{1\left(\widehat{d}\left(X_t\right)+\widehat{\overline{d}}\left(X_t\right)>\overline{\varepsilon}_t>0\right)-1\left(\widehat{d}\left(X_t\right)+\widehat{\overline{d}}\left(X_t\right)<\overline{\varepsilon}_t<0\right)\right\}w\left(X_t\right)\\
\equiv&A_T+B_T+C_T.
\end{aligned}$$

We first deal with the term B_T. From Lemma 2, we have

$$\begin{aligned}
B_T=&-\frac{1}{T(T-1)}\sum_{t\neq s}w\left(X_t\right)e_1'\frac{H_T^{-1}}{h^d}S_{T,p}^{-1}(X_t)K_h(X_s-X_t)\mu(X_s-X_t)\varphi(\overline{\varepsilon}_t)\varphi(\overline{\varepsilon}_s)+o_p\left(T^{-1/2}\right)\\
:=&B_{1T}+o_p\left(T^{-1/2}\right).
\end{aligned}$$

Following similar arguments to those used in the proof of Lemma 3 of Noh et al. (2013), one can prove that $B_{1T}=o_p\left(T^{-1/2}\right)$.

Next, by following the same steps for proving the asymptotic negligibility of B_{1T} and combining the asymptotic Bahadur representation for the nonparametric estimator $\widehat{m}\left(\cdot\right)$ - for instance the Nadaraya-Watson kernel estimator $\widehat{m}^{NW}\left(\cdot\right)$- we have

$$\widehat{m}^{NW}\left(x\right)-m\left(x\right)=b^2Bias_T\left(x\right)+\frac{1}{E\left[\widehat{f}_X\left(x\right)\right]}\frac{1}{T}\sum_{t=1}^{T}K_b\left(X_t-x\right)\sigma\left(X_t\right)\epsilon_t+R_T^*\left(x\right),$$

where the error terms ϵ_t are defined in (1), $Bias_T(x)$ is a deterministic bias term, and $\sup_{x\in\mathcal{D}}|R_T^*(x)| = O_p\left((Tb^d)^{-1}\log T\right) = o_p\left(T^{-1/2}\right)$. Thus, with the appropriate choice of the bandwidth b in Assumption **C.10**, we can show that $A_T = o_p\left(T^{-1/2}\right)$.

Finally, we deal with the term C_T. We first define $I(w) = \{t : X_t \in \mathcal{D}, t = 1, \dots, T\}$. As for the term C_T, notice that

$$
\begin{aligned}
|C_T| &\leq \frac{2}{T}\sum_{t=1}^{T}\left(|\overline{\varepsilon}_t| + \left|\widehat{d}(X_t) + \widehat{\overline{d}}(X_t)\right|\right)1\left(|\overline{\varepsilon}_t| < \left|\widehat{d}(X_t) + \widehat{\overline{d}}(X_t)\right|\right)w(X_t)\\
&\leq \frac{2}{T}\sum_{t=1}^{T}\left(|\overline{\varepsilon}_t| + \left|\widehat{d}(X_t)\right| + \left|\widehat{\overline{d}}(X_t)\right|\right)1\left(|\overline{\varepsilon}_t| < \left|\widehat{d}(X_t)\right| + \left|\widehat{\overline{d}}(X_t)\right|\right)w(X_t)\\
&\leq \frac{4}{T}\sum_{t=1}^{T}\left(\left|\widehat{d}(X_t)\right| + \left|\widehat{\overline{d}}(X_t)\right|\right)1\left(|\overline{\varepsilon}_t| < \left|\widehat{d}(X_t)\right| + \left|\widehat{\overline{d}}(X_t)\right|\right)w(X_t)\\
&\leq 4\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\max_{x\in\mathcal{D}}w(x)\,\frac{1}{T}\sum_{t=1}^{T}1\left(|\overline{\varepsilon}_t| < \max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\right).
\end{aligned}
$$

Using the Glivenko-Cantelli Theorem for strictly stationary sequences, we have

$$
\begin{aligned}
|C_T| \leq & 4\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\max_{x\in\mathcal{D}}w(x)\left\{Pr\left(|\overline{\varepsilon}| < \max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\right) + O_p\left(T^{-1/2}\right)\right\}\\
= & 4\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\max_{x\in\mathcal{D}}w(x)\left\{F_{\overline{\varepsilon}}\left(\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\right)\right.\\
& \left.-F_{\overline{\varepsilon}}\left(-\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\right)\right\} + 4\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\max_{x\in\mathcal{D}}w(x)\times O_p\left(T^{-1/2}\right)\\
\leq & 8\sup_{e\in\mathcal{E}}f_{\overline{\varepsilon}}(e)\max_{x\in\mathcal{D}}w(x)\left\{\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\right\}^2\\
& + 4\max_{s\in I(w)}\left(\left|\widehat{d}(X_s)\right| + \left|\widehat{\overline{d}}(X_s)\right|\right)\max_{x\in\mathcal{D}}w(x)\times O_p\left(T^{-1/2}\right),
\end{aligned}
$$

where the third step follows from the first order Taylor expansion of $F_{\overline{\varepsilon}}(e)$, with $F_{\overline{\varepsilon}}$ the distribution function of $\overline{\varepsilon}$ and $f_{\overline{\varepsilon}}$ the corresponding density function. Now, from the uniformly bounded density $f_{\overline{\varepsilon}}$, uniformly bounded weight function $w(x)$ over the support $\mathcal{D}$ in Assumption **C.5.**, and

$$
\max_{t\in I(w)}\left|\widehat{d}(X_t)\right| = O_p\left(\frac{\log T}{Tb^d}\right)^{3/4} = o_p\left(\frac{\log T}{Th^d}\right)^{3/4},
$$

under $h = o(b)$ in Assumption **C.10.**, and

$$
\max_{t\in I(w)}\left|\widehat{\overline{d}}(X_t)\right| = O_p\left(\frac{\log T}{Th^d}\right)^{3/4},
$$

from Kong et al. (2010) [Theorem 1, p.1536], it follows that

$$
C_T = O_p\left(\left(\frac{\log T}{Th^d}\right)^{3/2} + T^{-1/2}\left(\frac{\log T}{Th^d}\right)^{3/4}\right) = o_p\left(T^{-1/2}\right)
$$

under Assumption **C.10**. Combing the asymptotic negligibility of A_T, B_T and C_T proves Lemma 4. ■

Proof of Theorem 1: The proof of Theorem 1 is similar to but slightly simpler than that of Theorem 3 and therefore it is omitted.■

Proof of Theorem 2: The proof of Theorem 2 is similar to but slightly simpler than that of Theorem 4 and therefore it is omitted.■

Proof of Theorem 3: Theorem 3 can be proven using the two asymptotic representations in Lemmas 3 and 4 and the equality $\widehat{a}/\widehat{b} = a/b + \widehat{b}^{-1}\left[(\widehat{a}-a) - \left(\widehat{b}-b\right)(a/b)\right]$. ■

Proof of Theorem 4: Under the null hypothesis of homoskedasticity, the first order asymptotic result in Theorem 3 is degenerated. To establish the null limiting distribution of $\hat{H}(\widehat{m},\tau)$, we need to investigate the higher order terms of an analogous decomposition in Lemma 4. To this end, recall that for x and y, $\rho_\tau(x-y) - \rho_\tau(x) = (-y)\varphi(x) + 2(y-x)[1(y>x>0) - 1(y<x<0)]$. Denote $\varepsilon_t = Y_t - m(X_t) - \overline{\xi}(\tau)$ and $\overline{\varepsilon}_t = Y_t - m(X_t) - \overline{\phi}(X_t,\tau)$, respectively the restricted (under the null) and the unrestricted (under the alternative) quantile errors. Note that under the null hypothesis, $\varepsilon_t = \overline{\varepsilon}_t$ almost surely (a.s.) and we obtain

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\xi}}(\tau)\right)w(X_t) - \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\phi}}_{-t}(X_t,\tau)\right)w(X_t)\\
=&\left\{\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - m(X_t) - \overline{\xi}(\tau)\right)w(X_t) - \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{d}(X_t) + \widehat{\overline{d}}\right)w(X_t)\varphi(\varepsilon_t)\right.\\
&\left.-\frac{2}{T}\sum_{t=1}^{T}\left\{\varepsilon_t - \left[\widehat{d}(X_t) + \widehat{\overline{d}}\right]\right\}\left\{1\left(\widehat{d}(X_t) + \widehat{\overline{d}} > \varepsilon_t > 0\right) - 1\left(\widehat{d}(X_t) + \widehat{\overline{d}} < \varepsilon_t < 0\right)\right\}\right\}w(X_t)\\
&-\left\{\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - m(X_t) - \overline{\phi}(X_t,\tau)\right)w(X_t) - \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{d}(X_t) + \widehat{\overline{d}}(X_t)\right)w(X_t)\varphi(\overline{\varepsilon}_t)\right.\\
&\left.-\frac{2}{T}\sum_{t=1}^{T}\left\{\overline{\varepsilon}_t - \left[\widehat{d}(X_t) + \widehat{\overline{d}}(X_t)\right]\right\}\left\{1\left(\widehat{d}(X_t) + \widehat{\overline{d}}(X_t) > \overline{\varepsilon}_t > 0\right) - 1\left(\widehat{d}(X_t) + \widehat{\overline{d}}(X_t) < \overline{\varepsilon}_t < 0\right)\right\}\right\}w(X_t)\\
=&\frac{1}{T}\sum_{t=1}^{T}\widehat{\overline{d}}(X_t)w(X_t)\varphi(\overline{\varepsilon}_t) - \widehat{\overline{d}}\frac{1}{T}\sum_{t=1}^{T}w(X_t)\varphi(\varepsilon_t)\\
&+\frac{2}{T}\sum_{t=1}^{T}\left\{\overline{\varepsilon}_t - \left[\widehat{d}(X_t) + \widehat{\overline{d}}(X_t)\right]\right\}\left\{1\left(\widehat{d}(X_t) + \widehat{\overline{d}}(X_t) > \overline{\varepsilon}_t > 0\right) - 1\left(\widehat{d}(X_t) + \widehat{\overline{d}}(X_t) < \overline{\varepsilon}_t < 0\right)\right\}w(X_t)\\
&-\frac{2}{T}\sum_{t=1}^{T}\left\{\varepsilon_t - \left[\widehat{d}(X_t) + \widehat{\overline{d}}\right]\right\}\left\{1\left(\widehat{d}(X_t) + \widehat{\overline{d}} > \varepsilon_t > 0\right) - 1\left(\widehat{d}(X_t) + \widehat{\overline{d}} < \varepsilon_t < 0\right)\right\}w(X_t)\\
\equiv&D_T + E_T + F_T + G_T,
\end{aligned}
$$

with $\widehat{d}(x) = \widehat{m}(x) - m(x)$, $\widehat{\overline{d}} = \widehat{\overline{\xi}}(\tau) - \overline{\xi}(\tau)$, and $\widehat{\overline{d}}(x) = \widehat{\overline{\phi}}(x,\tau) - \overline{\phi}(x,\tau)$, where the second equality follows using $\varepsilon_t = \overline{\varepsilon}_t$ a.s.

First of all, recall that $F_T = C_T = O_p\left(\left(\frac{\log T}{Th^d}\right)^{3/2} + T^{-1/2}\left(\frac{\log T}{Th^d}\right)^{3/4}\right) = o_p\left(\left(Th^{d/2}\right)^{-1}\right)$. Following similar steps, we can also show that $G_T = O_p\left(\left(\frac{\log T}{Th^d}\right)^{3/2} + T^{-1/2}\left(\frac{\log T}{Th^d}\right)^{3/4}\right) = o_p\left(\left(Th^{d/2}\right)^{-1}\right)$.

In the subsequent analysis, we show that (i) $Th^{d/2}D_T$ converges in distribution to a zero mean normal variable with proper asymptotic variance; (ii) $E_T = o_p\left(\left(Th^{d/2}\right)^{-1}\right)$ under our assumptions in Appendix A.2.1. We start by proving point (ii). Observing that

$$E_T = \frac{1}{T}\left\{\sqrt{T}\widehat{\widehat{d}}\right\} \times \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} w\left(X_t\right)\varphi\left(\varepsilon_t\right)\right\} = \frac{1}{T} \times O_p(1) \times O_p(1) = O_p\left(\frac{1}{T}\right),$$

where $\sqrt{T}\widehat{\widehat{d}} = \sqrt{T}\left[\widehat{\widehat{\xi}}\left(\tau\right) - \overline{\xi}\left(\tau\right)\right] = O_p(1)$ follows immediately from the Bahadur representation in Lemma 1. On the other hand, the asymptotic normality of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} w\left(X_t\right)\varphi\left(\varepsilon_t\right)$ with zero mean follows from Assumptions **C.3** and **C.5**. and the null hypothesis which implies that $E[\varphi(\varepsilon_t)|X_t] = 0$. Thus, $E_T = O_p\left(1/T\right) = o_p\left(\left(Th^{d/2}\right)^{-1}\right)$. Now, we need to prove point (i). Let $f_X\left(x\right)$ and $f_{\overline{\varepsilon}|X}\left(0|x\right)$ denote the marginal density of X_t and the conditional density of $\overline{\varepsilon}$ given X_t and evaluated at $\overline{\varepsilon} = 0$, respectively. By Lemma 2, we have

$$\begin{aligned}
D_T =& -\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T} w\left(X_t\right)e_1'\frac{H_T^{-1}}{h^d}S_{T,p}^{-1}\left(X_t\right)K_h\left(X_s - X_t\right)\mu\left(X_s - X_t\right)\varphi(\overline{\varepsilon}_t)\varphi(\overline{\varepsilon}_s) + o_p\left(\left(Th^{d/2}\right)^{-1}\right)\\
=&\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\frac{w\left(X_t\right)}{f_{\overline{\varepsilon}|X}\left(0|X_t\right)f_X\left(X_t\right)}K_h\left(X_t - X_s\right)\varphi(\overline{\varepsilon}_t)\varphi(\overline{\varepsilon}_s) + o_p\left(\left(Th^{d/2}\right)^{-1}\right)\\
\equiv& D_{1T} + o_p\left(\left(Th^{d/2}\right)^{-1}\right),
\end{aligned}$$

where the second equality follows by applying the notion of "equivalent kernel" representation for local polynomial estimator [see Fan and Gijbels, 1996, pp.63-64],

We can now rewrite $Th^{d/2}D_{1T}$ into a classical U-statistic form with a symmetrized varying kernel which depends on the sample size T through the bandwidth h, i.e.

$$Th^{d/2}D_{1T} = \frac{2}{T-1}\sum_{1\leq t<s\leq T} U_T(\chi_t, \chi_s),$$

where $\chi_t = (X_t, \overline{\varepsilon}_t)$, $U_T(\chi_t, \chi_s) = \eta_T(\chi_t, \chi_s) + \eta_T(\chi_s, \chi_t)$, and

$$\eta_T(\chi_t, \chi_s) = \frac{w(X_t)}{2f_{\overline{\varepsilon}|X}(0|X_t)f_X(X_t)}\frac{1}{h^{\frac{d}{2}}}K\left(\frac{X_t - X_s}{h}\right)\varphi(\overline{\varepsilon}_t)\varphi(\overline{\varepsilon}_s).$$

Note that, under Assumption **C.3**., we have $E[U_T(\chi_t, \chi_s)] = E[\eta_T(\chi_t, \chi_s)] = E[U_T(\chi_t, \chi_s)|\chi_t] = E[\eta_T(\chi_t, \chi_s)|\chi_t] = 0$. Thus, the latter U-statistic is in fact a degenerated second order U-statistic. Under our Assumptions **C.1**, **C.3**, **C.6**, and **C.9**, one can check that the conditions of Theorem A.1 in Gao (2007) for second order degenerated U-statistic with strongly mixing processes are satisfied for the previous kernel $U_T(\chi_t, \chi_s)$ so that we can establish a central limit theory for the term $Th^{d/2}D_{1T}$ with asymptotic variance given by

$$\begin{aligned}
\tilde{\sigma}_{0\tau}^2 = \lim_{T\to\infty} 2E_tE_s\left[U_T^2(\chi_t, \chi_s)\right] &= \lim_{T\to\infty} 2E_tE_s\left[\eta_T(\chi_t, \chi_s)^2 + \eta_T(\chi_s, \chi_t)^2 + 2\eta_T(\chi_t, \chi_s)\eta_T(\chi_s, \chi_t)\right]\\
&= 2\tau^2\left(1-\tau\right)^2\int K^2(u)\,du\int\frac{w^2(x)}{f_{\overline{\varepsilon}|X}^2(0|x)}\,dx := \overline{r}_\tau^2\overline{\sigma}_{0\tau}^2,
\end{aligned}$$

where E_t denotes the expectation with respect to χ_t, and $\overline{r}_\tau = E\left[\rho_\tau\left(Y_t - m(X_t) - \overline{\xi}(\tau)\right)\right]$. The previous expression of asymptotic variance follows from a straightforward calculation of conditional expectation by combining integration, standard techniques of change of variables with Assumptions **C.7** and **C.9**. For instance, to calculate the first conditional expectation, it follows that

$$\lim_{T\to\infty} E_t E_s\left[\eta_T(\chi_t,\chi_s)^2\right] = \frac{1}{4}\tau^2(1-\tau)^2 \lim_{T\to\infty}\int\int \frac{w^2(x_1)}{f^2_{\overline{\varepsilon}|X}(0|x_1) f^2_X(x_1)} \frac{1}{h^d} K^2\left(\frac{x_1-x_2}{h}\right) f_X(x_1) f_X(x_2)\, dx_1\, dx_2$$
$$= \frac{1}{4}\tau^2(1-\tau)^2 \int K^2(u)\, du \int \frac{w^2(x)}{f^2_{\overline{\varepsilon}|X}(0|x)}\, dx.$$

Therefore, the CLT for the U-statistic form $Th^{d/2}D_{1T}$ together with the expression of asymptotic variance $\tilde{\sigma}^2_{0\tau}$ shows that $Th^{d/2}D_T = Th^{d/2}D_{1T} + o_p(1) \xrightarrow{d} N\left(0, \tilde{\sigma}^2_{0\tau}\right)$.

Finally, we have

$$\begin{aligned} Th^{d/2}\hat{H}^w(\widehat{m},\tau) &= \frac{1}{T^{-1}\sum_{t=1}^T \rho_\tau\left(Y_t - \widehat{m}(X_t) - \widehat{\overline{\xi}}(\tau)\right) w(X_t)}\left[Th^{d/2}D_T + o_p(1)\right] \\ &\xrightarrow{d} N\left(0, \overline{\sigma}^2_{0\tau}\right), \end{aligned}$$

where $\overline{\sigma}^2_{0\tau} = \tilde{\sigma}^2_{0\tau}/\overline{r}^2_\tau$. The last step follows naturally from Lemma 3 and the Slutsky's Theorem. Observe that a consistent estimator of $\tilde{\sigma}^2_{0\tau}$ is given by

$$\widehat{\tilde{\sigma}}^2_{0\tau} = 2\tau^2(1-\tau)^2 \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{s=1, s\neq t}^T \frac{w^2(X_t)}{\widehat{f}^2_{\overline{\varepsilon}|X}(0|X_t)\widehat{f}^2_X(X_t)} \frac{1}{h^d} K^2\left(\frac{X_t - X_s}{h}\right).$$

Like the term D_{1T}, the estimator $\widehat{\tilde{\sigma}}^2_{0\tau}$ can also be re-expressed as a U-statistic form

$$\widehat{\tilde{\sigma}}^2_{0\tau} = \frac{2}{T(T-1)} \sum_{1\leq t<s\leq T} H_T(X_t, X_s) + o_p(1),$$

with the following symmetrized kernel:

$$H_T(X_t, X_S) = \tau^2(1-\tau)^2\left(\frac{w^2(X_t)}{f^2_{\overline{\varepsilon}|X}(0|X_t) f^2_X(X_t)} + \frac{w^2(X_s)}{f^2_{\overline{\varepsilon}|X}(0|X_s) f^2_X(X_s)}\right)\frac{1}{h^d} K^2\left(\frac{X_t - X_s}{h}\right).$$

However, in contrast to D_{1T}, the second order U-statistic $\widehat{\tilde{\sigma}}^2_{0\tau}$ is a non-degenerated one. By applying a standard Hoeffding decomposition on the previous expression, one can show that $\widehat{\tilde{\sigma}}^2_{0\tau} = \tilde{\sigma}^2_{0\tau} + o_p(1)$. This concludes the proof of Theorem 4. ■

Proof of Proposition 1: Using the convergence (in probability) result implied by the Bahadur representation in Theorem 3, one can readily see that if the alternative hypothesis H_A of heteroskedasticity is true, then on the one hand we have: (i) $\hat{H}^w(\widehat{m},\tau) = H^w(m,\tau) + o_p(1)$, where $H(m,\tau) > 0$. On the other hand, following arguments similar to those we used in the proof of the consistency of the asymptotic variance estimator $\widehat{\tilde{\sigma}}^2_{0\tau}$ in Theorem 4, we can show that (ii) $\widehat{\overline{\sigma}}^2_{0\tau} = O_p(1)$ under heteroskedasticity. Proposition 1 follows then from (i) and (ii). ■

Proof of Theorem 5: It is similar to the proof of Theorem 4, thus we only provide a sketch. Define

$$H_T = \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\xi}\left(\tau\right)\right)w\left(X_t\right) - \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right).$$

Then, using the same steps as in the proof of Theorem 4, under H_{1T} we have the decomposition:

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}\left(X_t\right) - \widehat{\overline{\xi}}\left(\tau\right)\right)w\left(X_t\right) - \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}\left(X_t\right) - \widehat{\overline{\phi}}_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)\\ \equiv&D_T + E_T + F_T + G_T + H_T.\end{aligned}$$

Here, following the same arguments as in the proof of Theorem 4, $Th^{d/2}D_T \to N(0,\tilde{\sigma}^2_{0\tau})$ in distribution, $E_T = F_T = G_T = o_p\left(\left(Th^{d/2}\right)^{-1}\right)$. However, by the law of large numbers and the second order Taylor expansion of $E[\rho_\tau(\cdot)]$, the remaining term H_T becomes

$$\begin{aligned}H_T =&E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\xi}\left(\tau\right)\right)w\left(X_t\right)\right] - E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right] + o_p\left(1\right)\\ =&E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right) + \frac{D^{-1}\left(\tau\right)}{T^{1/2}h^{d/4}}\Delta_T\left(X_t\right)\right)w\left(X_t\right)\right]\\ &- E\left[\rho_\tau\left(Y_t - m\left(X_t\right) - \overline{\phi}\left(X_t,\tau\right)\right)w\left(X_t\right)\right] + o_p\left(1\right)\\ =&\frac{D^{-1}\left(\tau\right)}{T^{1/2}h^{d/4}}E\left[\Delta_T\left(X_t\right)w\left(X_t\right)\varphi(\overline{\varepsilon}_t)\right] - \left(\frac{D^{-1}\left(\tau\right)}{T^{1/2}h^{d/4}}\right)^2 E\left[\Delta_T^2\left(X_t\right)w\left(X_t\right)g\left(X_t\right)\right] + o_p\left(1\right)\\ =&\left(\frac{D^{-1}\left(\tau\right)}{T^{1/2}h^{d/4}}\right)^2 E\left[\Delta_T^2\left(X_t\right)w\left(X_t\right)f_{\overline{\varepsilon}|X}\left(0|X_t\right)\right] + o_p\left(1\right),\end{aligned}$$

where $g\left(x\right) = \partial E\left[\varphi\left(X_t-\theta\right)|X_t = x\right]/\partial\theta = -f_{\overline{\varepsilon}|X}\left(0|x\right)$, and the last step follows from the fact that $E\left[\varphi(\overline{\varepsilon}_t)|X_t\right] = 0$ and the law of iterated expectation. By the preceding information, Lemma 3 and the Slutsky's theorem, we have

$$\begin{aligned}Th^{d/2}\hat{H}^w\left(\widehat{m},\tau\right) =& \frac{Th^{d/2}D_T + Th^{d/2}H_T}{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(Y_t - \widehat{m}\left(X_t\right) - \widehat{\overline{\xi}}\left(\tau\right)\right)w\left(X_t\right)}\times\left[1+o_p\left(1\right)\right]\\ &= \overline{r}_\tau^{-1}\left\{Th^{d/2}D_T + \left(D^{-1}\left(\tau\right)\right)^2 E\left[\Delta_T^2\left(X_t\right)w\left(X_t\right)f_{\overline{\varepsilon}|X}\left(0|X_t\right)\right]\right\}\times\left[1+o_p\left(1\right)\right]\\ &\xrightarrow{d} N\left(\gamma,\overline{\sigma}^2_{0\tau}\right)\end{aligned}$$

where $\gamma := \overline{r}_\tau^{-1}\left(D^{-1}\left(\tau\right)\right)^2\lim_{T\to\infty}E\left[\Delta_T^2\left(X_t\right)w\left(X_t\right)f_{\overline{\varepsilon}|X}\left(0|X_t\right)\right] > 0$ is the non-zero mean term and $\overline{\sigma}^2_{0\tau} = \tilde{\sigma}^2_{0\tau}/\overline{r}^2_\tau$ the asymptotic variance. Therefore, we have shown that $Th^{d/2}\hat{H}^w\left(\widehat{m},\tau\right) \to N(\gamma,\overline{\sigma}^2_{0\tau})$ in distribution under the local alternatives H_{1T} in (16). ■

Proof of Theorem 6: Conditionally on $\left\{(X_t', Y_t)'\right\}_{t=1}^T$, Theorem 6 can be proved using similar arguments to the ones we used in the proof of Theorem 4. Analogously, let $\varepsilon_t^* = Y_t^* - \widehat{m}\left(X_t\right) - \overline{\xi}^*\left(\tau\right)$ and $\overline{\varepsilon}_t^* = Y_t^* - \widehat{m}\left(X_t\right) - \overline{\phi}^*\left(X_t,\tau\right)$, respectively, the restricted and the unrestricted quantile errors conditional on the original sample $\left\{(X_t', Y_t)'\right\}_{t=1}^T$. Let $E^*(\cdot)$ denote the expectation and $o_{p^*}(1)$ the convergence in probability under the bootstrap law. The notation $O_{p^*}(1)$ is similarly defined. We have

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t^*-\widehat{m}^*\left(X_t\right)-\widehat{\overline{\xi}}^*\left(\tau\right)\right)w\left(X_t\right)-\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_t^*-\widehat{m}^*\left(X_t\right)-\widehat{\overline{\phi}}^*_{-t}\left(X_t,\tau\right)\right)w\left(X_t\right)\\
=&\frac{1}{T}\sum_{t=1}^{T}\widehat{\overline{d}}^*\left(X_t\right)w\left(X_t\right)\varphi\left(\overline{\varepsilon}_t^*\right)-\widehat{\overline{d}}^*\frac{1}{T}\sum_{t=1}^{T}w\left(X_t\right)\varphi\left(\varepsilon_t^*\right)\\
&+\frac{2}{T}\sum_{t=1}^{T}\left\{\overline{\varepsilon}_t^*-\left[\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*\left(X_t\right)\right]\right\}\left\{1\left(\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*\left(X_t\right)>\overline{\varepsilon}_t^*>0\right)-1\left(\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*\left(X_t\right)<\overline{\varepsilon}_t^*<0\right)\right\}w\left(X_t\right)\\
&-\frac{2}{T}\sum_{t=1}^{T}\left\{\varepsilon_t^*-\left[\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*\right]\right\}\left\{1\left(\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*>\varepsilon_t^*>0\right)-1\left(\widehat{d}^*\left(X_t\right)+\widehat{\overline{d}}^*<\varepsilon_t^*<0\right)\right\}w\left(X_t\right)\\
\equiv&D_T^*+E_T^*+F_T^*+G_T^*,
\end{aligned}
$$

where D_T^*, E_T^*, F_T^* and G_T^* are the bootstrap analogue of D_T, E_T, F_T and G_T, and $\widehat{d}^*\left(x\right)=\widehat{m}^*\left(x\right)-\widehat{m}\left(x\right)$, $\widehat{\overline{d}}^*=\widehat{\overline{\xi}}^*\left(\tau\right)-\overline{\xi}^*\left(\tau\right)$, and $\widehat{\overline{d}}^*\left(x\right)=\widehat{\overline{\phi}}^*\left(x,\tau\right)-\overline{\phi}^*\left(x,\tau\right)$.

Noting that $\max_{t\in I(w)}\left|\widehat{d}^*\left(X_t\right)\right|=o_{p^*}\left(\log T/\left(Th^d\right)\right)^{3/4}$, $\max_{t\in I(w)}\left|\widehat{\overline{d}}^*\left(X_t\right)\right|=o_{p^*}\left(\log T/\left(Th^d\right)\right)^{3/4}$, and $\widehat{\overline{d}}^*=O_{p^*}\left(T^{-1/2}\right)$ under our assumptions, the proof of $E_T^*=o_{p^*}(1)$, $F_T^*=o_{p^*}(1)$ and $G_T^*=o_{p^*}(1)$ is analogous to that of E_T, F_T and G_T in the proof of Theorem 4 and is thus omitted.

We can show that

$$
\begin{aligned}
&Th^{d/2}D_T^*\\
=&\frac{h^{d/2}}{(T-1)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\frac{w\left(X_t\right)}{f_{\overline{\varepsilon}^*|X}\left(0|X_t\right)f_X\left(X_t\right)}K_h\left(X_t-X_s\right)\varphi(\overline{\varepsilon}_t^*)\varphi(\overline{\varepsilon}_s^*)+o_{p^*}\left(1\right)\\
=&\frac{1}{T-1}\sum_{1\le t<s\le T}\left(\frac{w(X_t)}{f_{\overline{\varepsilon}^*|X}(0|X_t)f_X(X_t)}\frac{1}{h^{\frac{d}{2}}}K\left(\frac{X_t-X_s}{h}\right)\varphi(\overline{\varepsilon}_t^*)\varphi(\overline{\varepsilon}_s^*)\right.\\
&\left.+\frac{w(X_s)}{f_{\overline{\varepsilon}^*|X}(0|X_s)f_X(X_s)}\frac{1}{h^{\frac{d}{2}}}K\left(\frac{X_t-X_s}{h}\right)\varphi(\overline{\varepsilon}_t^*)\varphi(\overline{\varepsilon}_s^*)\right)+o_{p^*}\left(1\right)\\
\equiv&Th^{d/2}D_{1T}^*+o_{p^*}\left(1\right).
\end{aligned}
$$

Noting that $Th^{d/2}D_{1T}^*$ is a second order degenerate U-statistic and $\overline{\varepsilon}_t^*$'s are i.i.d. with τ-th quantile 0 and satisfying $E^*(\overline{\varepsilon}_t^*|X_t)=0$, conditional on the original sample. We can apply the central limit theorem for second order degenerate U-statistic for i.i.d. observations (e.g., Hall, 1984) and conclude that conditional on the original sample, $Th^{d/2}D_{1T}^*\xrightarrow{d}N(0,\tilde{\sigma}_{0\tau}^{*2})$, where the asymptotic variance is

$$
\tilde{\sigma}_{0\tau}^{*2}=2\tau^2\left(1-\tau\right)^2\int K^2(u)\,du\int\frac{w^2(x)}{f^2_{\overline{\varepsilon}^*|X}(0|x)}\,dx,
$$

with a consistent estimator of $\tilde{\sigma}_{0\tau}^{*2}$ given by

$$
\widehat{\tilde{\sigma}}_{0\tau}^{*2}=2\tau^2\left(1-\tau\right)^2\frac{1}{T\left(T-1\right)}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\frac{w^2(X_t)}{\widehat{f}^2_{\overline{\varepsilon}^*|X}(0|X_t)\widehat{f}_X^2(X_t)}\frac{1}{h^d}K^2\left(\frac{X_t-X_s}{h}\right).
$$

Therefore, the bootstrapped measure $\widehat{H}^{w*}(\widehat{m}^*, \tau)$ satisfies

$$Th^{d/2}\widehat{H}^{w*}\left(\widehat{m}^*, \tau\right) = \frac{Th^{d/2}D_{1T}^*}{T^{-1}\sum_{t=1}^{T}\rho_\tau\left(Y_t^* - \widehat{m}^*\left(X_t\right) - \widehat{\overline{\xi}}^*\left(\tau\right)\right)w\left(X_t\right)} + o_{p^*}\left(1\right) \xrightarrow{d} N\left(0, \overline{\sigma}_{0\tau}^{*2}\right),$$

where $\overline{\sigma}_{0\tau}^{*2} = \tilde{\sigma}_{0\tau}^{*2}/\overline{r}_\tau^{*2}$ and $\overline{r}_\tau^* = E\left[\rho_\tau\left(Y_t^* - \widehat{m}\left(X_t\right) - \overline{\phi}^*\left(X_t, \tau\right)\right)w\left(X_t\right)\right]$. Finally, conditional on the data, we conclude that the bootstrapped test statistic $\widehat{\overline{\Gamma}}^* = Th^{d/2}\widehat{H}^{w*}\left(\widehat{m}^*, \tau\right)/\widehat{\overline{\sigma}}_{0\tau}^* \xrightarrow{d} N\left(0, 1\right)$, where $\widehat{\overline{\sigma}}_{0\tau}^{*2} = \widehat{\tilde{\sigma}}_{0\tau}^{*2}/\widehat{\overline{r}}_\tau^{*2}$ and $\widehat{\overline{r}}_\tau^*$ is the sample analogue of $\overline{r}_\tau^*$.